

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 ; Search time 60 Seconds
(without alignments)
1323.264 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500
Sequence: 1 MWSKGEBELFTGVPIVLELD.....AQESGMDRHPACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	100.0	281	3	AAV50142 Green flu
2	1500	100.0	281	3	AAV50142 Green flu
3	1500	100.0	281	3	AAV50142 Green flu
4	1289.5	86.0	805	3	AAU10888 EGFP-MODC
5	1289.5	86.0	805	3	AAU10888 EGFP-MODC
6	1280.5	85.4	793	5	ABG94503 Protease
7	1280.5	85.4	793	5	ABG94503 Protease
8	1279.5	85.3	357	6	ABP72573 GFP-Droso
9	1279.5	85.3	357	6	ABP72573 GFP-Droso
10	1278	85.2	607	2	AAW85032 Green flu
11	1278	85.2	607	2	AAW85032 Green flu
12	1277	85.1	1089	7	AAW85574 Hs-DNC-53
13	1276.5	85.1	284	7	ADH90657 Human and
14	1276.5	85.1	1090	7	ADH90657 Human and
15	1276	85.1	286	7	ADH90661 Human and
16	1276	85.1	289	7	ADH90661 Human and
17	1276	85.1	290	7	ADH90661 Human and
18	1276	85.1	290	7	ADH90661 Human and
19	1276	85.1	290	7	ADH90661 Human and
20	1276	85.1	290	7	ADH90661 Human and
21	1276	85.1	290	7	ADH90661 Human and
22	1276	85.1	290	7	ADH90661 Human and
23	1275.5	85.0	941	3	AAW85039 Green flu
24	1275.5	85.0	941	3	AAW85039 Green flu
25	1275.5	85.0	941	3	AAW85039 Green flu

26	1275	85.0	432	6	ABP56093 TAT-EGFP-
27	1275	85.0	727	2	AAW85041 Green flu
28	1275	85.0	797	2	AAW85043 NFKappaB
29	1275	85.0	797	3	AAV70782 EGFP-NFka
30	1275	85.0	918	3	AAV70782 EGFP-NFka
31	1274.5	85.0	255	4	AAW86143 A. victor
32	1274	84.9	239	3	AAW82882 Enhanced
33	1274	84.9	239	3	AAV54349 Amino aci
34	1274	84.9	239	3	AAV79584 EGFP sign
35	1274	84.9	239	4	AAW50804 Jellyfish
36	1274	84.9	239	4	AAW50804 A. victor
37	1274	84.9	239	4	AAW50804 A. victor
38	1274	84.9	239	5	ABG66198 A. victor
39	1274	84.9	239	5	ABG66198 A. victor
40	1274	84.9	239	5	ABG66198 A. victor
41	1274	84.9	239	6	AAE34958 Aequorea
42	1274	84.9	239	6	AAE34958 Aequorea
43	1274	84.9	239	6	ABR83616 Green flu
44	1274	84.9	239	6	ADA38074 Aequorea
45	1274	84.9	239	7	ABU63204 Aequorea

ALIGNMENTS

RESULT 1

AAV50142 standard; protein; 281 AA.

AAV50142;

12-SEP-2003 (revised)
31-JAN-2000 (first entry)

Green fluorescent protein/ornithine decarboxylase fusion gene.

Enhanced green fluorescent protein; EGFP; ornithine decarboxylase; MODC; fusion; reporter; degradation domain; C-terminal; gene expression; KW protein localization; fluorescence; short-lived; destabilized; half-life; KW rapid degradation; EST sequence; EGFP-MODC; EGFP-MODC422-461; KW rapid turnover; determination; short-term event; repetitive event; KW sensitive; developmental process; transient; periodic; cyclic; KW expression; protein transport; circadian rhythm; kinetic; dynamic; KW transcription.

Aequorea victoria.

Mus sp.

Chimeric.

WC0954348-A1.

28-OCT-1999.

13-NOV-1998; 98WO-US024323.

17-APR-1998; 98US-00062102.

(CLON-) CLONTECH LAB INC.

Li X, Kain S;

WPI; 2000-013225/01.

N-PSDB; AA232767.

New fusion protein useful for studying cell localization, for studying

cell lineage and for assaying activation or deactivation of

transcriptional or translational elements.

Claim 5; Page 23; 49p; English.

This sequence represents a novel fusion protein, EGFP-MODC422-461,

comprising an enhanced variant of green fluorescent protein (EGFP) fused

via its C-terminus to residues 422-461 of the C-terminal degradation

CC domain of mouse ornithine decarboxylase (MODC). Green fluorescent protein (GFP) is a widely used reporter in studies of gene expression and protein localization; however, it is a very stable protein which can therefore accumulate and thus is often toxic to mammalian cells. MODC, in contrast, is one of the most short-lived proteins in mammalian cells, with a half life of about 30 minutes. Its rapid degradation is due to the composition of its C-terminus, a portion of which contains a PST sequence. The PST sequence contains a region enriched with proline, glutamate, serine and threonine, often flanked by basic amino acids, and has been proposed as characterizing short-lived proteins. The EGFP-MODC fusion protein, which comprises EGFP and the C-terminal MODC degradation domain, including a PST sequence, combines the fluorescence of GFP with the short life of MODC, and has a fluorescence half-life of about 2 hours in vivo. As the new fusion protein is destabilised with a rapid turnover, it is useful for producing stable cell lines which express GFP without the toxicity caused by build-up of GFP. Unlike prior art GFPs, the new fusion protein is useful for determining short-term or repetitive events, and produces more sensitive results due to the lack of accumulation of GFP. GFP fluorescence can be detected without additional expensive steps, e.g., lysing cells, adding exogenous substrates. The destabilised GFP is especially useful for studying developmental processes where genes are transiently expressed, dynamics of protein transport, and periodic and cyclic expression of genes that control unique biological phenomena such as circadian rhythms, and is useful as a reporter to study the kinetics of mRNA transcription from a regulated promoter. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 281 AA;

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 6.7e-146;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MWSKGEELFTGVVPLIVELDGVNGHKPSVSGEGGATYGLTLKFTCTGKLPVMPPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKPSVSGEGGATYGLTLKFTCTGKLPVMPPT 60
QY 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYTRAEVFEFGPTL 120
DB 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYTRAEVFEFGPTL 120
QY 121 VNRLEKIDGDFEKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDGDFEKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGPGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGPGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281

```

RESULT 2

AAB24252
ID AAB24252 standard; protein; 281 AA.

AC AAB24252;

DT 08-FEB-2001 (first entry)

DE EGFP-MODC422-461 fusion protein SEQ ID NO:1.

XX Green fluorescent protein; GFP; jellyfish; EGFP; mouse; ODC; MODC;
XX ornithine decarboxylase; PEST; fusion protein; genetic reporter;
XX identification.

OS Aequorea victoria.

OS Mus musculus.

OS Synthetic.

OS US6130313-A.

XX 10-OCT-2000.
XX
XX
XX 17-APR-1998; 98US-00062102.
XX
XX 02-OCT-1997; 97US-0060855P.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Kain S, Li X;
XX
XX WPI; 2000-013225/01.
XX N-PSDB; AAC64428.
XX
XX New fusion protein useful for studying cell localization, for studying
XX cell lineage and for assaying activation or deactivation of
XX transcriptional or translational elements.
XX
XX Claim 5; Col 12; 18pp; English.

Sequence 281 AA;

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 6.7e-146;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MWSKGEELFTGVVPLIVELDGVNGHKPSVSGEGGATYGLTLKFTCTGKLPVMPPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKPSVSGEGGATYGLTLKFTCTGKLPVMPPT 60
QY 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYTRAEVFEFGPTL 120
DB 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYTRAEVFEFGPTL 120
QY 121 VNRLEKIDGDFEKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDGDFEKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGPGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGPGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281

```

RESULT 3

AAU10888
ID AAU10888 standard; protein; 281 AA.

AC AAU10888;

DT 29-AUG-2003 (revised)

DT 14-FEB-2002 (first entry)

DE EGFP-MODC422-461 fusion protein.

XX Green fluorescent protein; ornithine decarboxylase; EGFP/MODC422-461;
XX PEST region; protein degradation; drug screening;
XX transcriptional control sequence; translational control sequence.

OS Aequorea victoria.

OS Mus sp.

OS Chimeric.

```

XX Key Location/Qualifiers
FH 1..242
FT Protein /label= Green fluorescent protein
FT Peptide 243..281
FT /label= MOC422-461
FT /note= "Murine ornithine decarboxylase residues 442-461"
FT 243..270
FT /label= PEST_region
FT /note= "Region enriched with Pro, Glu, Ser and Thr,
FT proposed to confer a short half-life on the protein"
XX
XX US6306600-B1.
XX
XX 23-OCT-2001.
XX
XX 30-JUL-1999; 99US-00364946.
XX
XX 17-APR-1998; 98US-00062102.
XX 13-NOV-1998; 98US-00191233.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Kain S, Li X;
XX
XX MPI; 2002-040198/05.
XX N-PSDB; AAS16105.
XX
XX Assaying the regulatory function of a transcriptional control sequence,
XX for drug screening or for analyzing protein degradation, comprises
XX employing rapidly degrading Green Fluorescent Protein-fusion proteins.
XX
XX Example 6; Col 12; 25pp; English.
XX
XX The invention relates to assaying the regulatory function of a
XX transcriptional or translational control sequence, with a transient
XX fluorescent reporter protein, comprises transfecting cells with an
XX expression vector having a DNA sequence that codes for a fluorescent
XX fusion protein comprising a Green Fluorescent Protein (GFP) and a PEST (a
XX region enriched for Pro, Glu, Ser and Thr residues) sequence. The fusion
XX protein, due to the presence of the PEST sequence has a half-life of no
XX ore than 10 hours. The method is useful for analysing transcriptional
XX regulation and/or cis-acting regulatory elements. The method is also
XX useful for studying protein degradation. The method is also useful for
XX studying processes involving multiple gene expression, as well as drug
XX screening assays. The present method permits easier development of stable
XX cell lines that express the GFP gene, since toxic levels of GFP are
XX avoided because the GFP protein is degraded quickly. Furthermore, the
XX destabilised EGFP-MODC (a humanised GFP/murine ornithine decarboxylase
XX PEST region fusion protein) decreases GFP accumulation. Accumulation of
XX fluorescent protein can interfere with the sensitivity of analysis. Thus,
XX the destabilised, rapid turnover fusion protein renders more sensitive
XX results. The present sequence is EGFP-MOC422-461, a humanised GFP/murine
XX ornithine decarboxylase PEST region fusion protein for use in the method
XX of the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 281 AA;
XX
XX Query Match 100.0%; Score 1500; DB 5; Length 281;
XX Best Local Similarity 100.0%; Pred. No. 6,7e-146;
XX Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 181 DRYQANTPIGDGPVLTLPDNNHLYSTQSALSKDPNKKRPHWLTBPTAAGTITLGMDELKXK 240
Db 181 DRYQANTPIGDGPVLTLPDNNHLYSTQSALSKDPNKKRPHWLTBPTAAGTITLGMDELKXK 240
Oy 241 LSHGFPEVEERODDGTLPMSCAQESGMDRHPACASARINV 281
Db 241 LSHGFPEVEERQDDGTLPMSCAQESGMDRHPACASARINV 281
RESULT 4
AAB22940
ID AAB22940 standard; protein; 805 AA.
XX
XX AAB22940;
XX
XX 10-JAN-2001 (first entry)
XX
XX GFP-NF-kappa-B fusion protein, SEQ ID NO:178.
XX
XX Biotector protein; fusion protein; recognition site;
XX cellular targeting sequence; cellular localisation; fluorescent protein;
XX protease activity detection; toxin detection; cellular stress detection;
XX drug discovery; cell based screening.
XX
XX Aegorrea victoria.
XX Unidentified.
XX Chimeric.
XX
XX WO200050872-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US004794.
XX
XX 26-FEB-1999; 99US-0122152P.
XX 08-MAR-1999; 99US-0123399P.
XX 12-JUL-1999; 99US-00352171.
XX
XX (CELL-) CELLOMICS INC.
XX
XX Giuliano KA, Kapur R;
XX
XX MPI; 2000-594086/56.
XX N-PSDB; AAA93445.
XX
XX Automated cell-based characterization of toxin by contacting cells
XX containing luminescent reporter molecules with test substance and
XX analyzing optically.
XX
XX Example 11; Page 329-332; 336pp; English.
XX
XX The invention relates to systems, methods and reagents for cell-based
XX screening or detection of compounds which affect particular biological
XX functions. The methods of the invention utilise fluorescent biodeceptor
XX molecules which, when acted on by a compound of interest, cause an
XX alteration in the cellular distribution of at least the fluorescent
XX moiety. In one embodiment, the biosensors comprise heat shock proteins
XX (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent
XX protein (GFP), or derivatives thereof). Such biosensors are located in
XX the cytoplasm, but on stress activation translocate to the nucleus. In
XX another embodiment biodeceptor proteins can be used to detect protease
XX activity. Such protease biodeceptor fusion proteins comprise one or more
XX fluorescent proteins; a recognition signal which is cleaved by the
XX protease; and at least one cellular localisation signal. The latter two
XX components may be components of a single protein which is acted upon by
XX the protease, or may be from heterologous sources. Due to the
XX localisation signal, the biodeceptor protein is localised to a particular
XX region of the cell. Once acted on by the protease of interest, the
XX fluorescent protein is cleaved from the localisation sequence, and is
XX free to migrate to other locations within the cell. The presence of a
XX second localisation signal attached to the fluorescent protein enables
XX the fluorescent protein to be directed to a different cellular

```


CC compartment after cleavage of the protease recognition sequence. The
CC change in distribution of the fluorescent protein can be detected using
CC imaging methods with a high degree of spatial resolution. The methods and
CC biosensors of the invention can be used to investigate a wide range of
CC cellular activities and to screen compounds which modulate these
CC activities. Biosensors containing a recognition site for caspase, for
CC example, may be used for the screening of compounds which modulate
CC apoptosis, while biosensors containing other protease recognition sites
CC may be used for the detection of proteolytic toxins (such as anthrax
CC lethal factor). The method provides improved target validation and
CC candidate compound optimization by combining many cell screening formats
CC with fluorescence-based molecular reagents and computer-based feature
CC extraction, data analysis and automation, resulting in increased quantity
CC and speed of data collection and faster evaluation of drug candidates.
CC Sequences AAB22860-B22876 and AAB22936-B22941 represent biosensor fusion
CC proteins produced in an exemplification of the invention
CC
SQ Sequence 805 AA;

Query Match 86.0%; Score 1289.5; DB 3; Length 805;
Best Local Similarity 92.9%; Pred. No. 1.7e-123;
Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIVLVDGVDVNGHGFSGSGEGDATTGKTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLVDGVDVNGHGFSGSGEGDATTGKTLKFTCTTGKLPVPMPT 60
QY 61 LVTTITLVGVCFSRYPDPMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAEKFGDPTL 120
DB 61 LVTTITLVGVCFSRYPDPMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAEKFGDPTL 120
QY 121 VNRTELKGDIDFKEDGNILGHKLEYNVNSHNYITMADKQKNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRTELKGDIDFKEDGNILGHKLEYNVNSHNYITMADKQKNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDGPVLLPNNHVLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDGPVLLPNNHVLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
QY 241 -LSHGFPPEVEEODDGTLPMSCAQESG 266
DB 241 GLRSRDPPEMDELFPPIFPAEPAQASG 267

RESULT 5

ABG94503
ID ABG94503 standard; protein; 805 AA.

AC ABG94503;

DT 27-NOV-2002 (first entry)

DE Protease biosensor related biosensor #5.

KM Detection; classification; identification; toxin detection; protease;
KM ADP-ribosylating toxin; cytotoxic phospholipase; exfoliative toxin;
KM toxic threat agent.

OS Synthetic.

XX US6416959-B1.

PD 09-JUL-2002.

PF 25-FEB-2000; 2000US-00513783.

XX 27-FEB-1997; 97US-00810983.
XX 27-FEB-1998; 98US-00031271.
XX 26-FEB-1999; 99US-0122152P.
XX 08-MAR-1999; 99US-0123399P.
XX 12-JUL-1999; 99US-00352171.
XX 31-AUG-1999; 99US-0151797P.
XX 17-SEP-1999; 99US-00398965.

PR 29-OCT-1999; 99US-00430656.
PR 01-DEC-1999; 99US-0168408P.
XX
PA (GITU/) GIULIANO K.
PA (KAPU/) KAPUR R.
XX
PI Giuliano K, Kapur R;
XX
XX WPI; 2002-634730/68.
DR N-PSDB; ABB871564.

PT Automated cell-based toxin detection, classification, and/or
PT identification by treating cells involves use of three classes of
PT luminescent reporter molecules such as detectors, classifiers or
PT identifiers.

PS Example 11; Col 311-316; 214pp; English.

CC The invention describes methods of automated detection, classification
CC and identification comprising treating cells containing luminescent
CC reporter molecules (I) in array of locations with a test substance, where
CC (I) are detectors, classifiers or identifiers, imaging cells in each
CC location to obtain luminescent signals and converting optical information
CC into digital data to interpret presence of toxins in the test substance.
CC The method are useful for detection of toxins chosen from proteases, ADP-
CC ribosylating toxins, cytotoxic phospholipases, and exfoliative toxins.
CC Three classes of cell-based luminescent reporter molecules such as
CC detectors, classifiers and identifiers are described and serve as
CC reporters of toxic threat agents. The first two levels of
CC characterisation ensure a rapid readout of toxin class without
CC sacrificing the ability to detect many new mutant toxins or dissect
CC several complex mixtures of known toxins. This is the amino acid sequence
CC of a biosensor associated with the protease biosensor of the invention
CC
SQ Sequence 805 AA;

Query Match 86.0%; Score 1289.5; DB 5; Length 805;
Best Local Similarity 92.9%; Pred. No. 1.7e-123;
Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIVLVDGVDVNGHGFSGSGEGDATTGKTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLVDGVDVNGHGFSGSGEGDATTGKTLKFTCTTGKLPVPMPT 60
QY 61 LVTTITLVGVCFSRYPDPMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAEKFGDPTL 120
DB 61 LVTTITLVGVCFSRYPDPMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAEKFGDPTL 120
QY 121 VNRTELKGDIDFKEDGNILGHKLEYNVNSHNYITMADKQKNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRTELKGDIDFKEDGNILGHKLEYNVNSHNYITMADKQKNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDGPVLLPNNHVLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDGPVLLPNNHVLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
QY 241 -LSHGFPPEVEEODDGTLPMSCAQESG 266
DB 241 GLRSRDPPEMDELFPPIFPAEPAQASG 267

RESULT 6

ABB80756
ID ABB80756 standard; protein; 793 AA.

AC ABB80756;

DT 23-SEP-2002 (first entry)

DE Amino acid sequence of a fusion GFP/HER beta protein.

KM Fluorometric analyzer; laser; fluorescence; GFP; ER beta; human; gene;
KM green fluorescent protein; estrogen receptor beta; fusion protein.

XX OS Synthetic.
 XX OS Aequorea victoria.
 XX OS Homo sapiens.
 XX PN WO200248693-A1.
 XX PD 20-JUN-2002.
 XX PF 14-DEC-2001; 2001WO-0P010998.
 XX PF 14-DEC-2000; 2000JP-00380327.
 XX PR 30-JAN-2001; 2001JP-00022105.
 XX PA (OLYU) OLYMPUS OPTICAL CO LTD.
 XX PI Kato N, Sakamoto H;
 XX DR WPI; 2002-528176/56.
 XX DR N-PSDB; AAN86125.
 XX PS Example 2; Page 98-105; 1093p; Japanese.
 XX CC The invention relates to a fluorometric analyzer that comprises a laser
 CC light source, an optical system for focusing the light beam from laser
 CC light source to a sample to form a confocal region, another optical
 CC system for focusing fluorescence from the sample, a light detector for
 CC measuring intensity, and a recording means to record the variation of
 CC intensity in the measured fluorescence. The analyzers are for studying
 CC behaviors of fluorescence-labeled molecules particularly intracellular
 CC biological molecules like proteins at molecular level, e.g. protein
 CC functions and interactions. Such analyzers are stable and convenient to
 CC operate, thereby enabling easy performance of fluorescence correlation
 CC spectroscopy, fluorescence intensity distribution analysis and
 CC fluorescence intensity multiple distribution analysis. The present
 CC sequence represents the amino acid sequence of a fusion protein
 CC comprising a GFP (green fluorescent protein) and hER (human estrogen
 CC receptor) beta protein
 CC XX
 XX SQ Sequence 793 AA;
 XX Query Match 85.4%; Score 1280.5; DB 5; Length 793;
 XX Best Local Similarity 88.4%; Pred. No. 1.4e-122;
 XX Matches 245; Conservative 2; Mismatches 5; Indels 25; Gaps 1;
 QY 1 MWSKGEELFTGVVPIIVELDGVNKHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 Db 1 MWSKGEELFTGVVPIIVELDGVNKHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 QY 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDYL 120
 Db 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDYL 120
 QY 121 VNRTELKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 Db 121 VNRTELKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 QY 241 LSHRFPPEVVEQDDGTLPMSCAQSSGMDRHPAAQASA 277
 Db 240 -----SGLRSRAQACASS 252

ID AAN85037 standard; protein; 1171 AA.
 XX AAN85037;
 AC 17-OCT-2003 (revised)
 XX 08-FEB-1999 (first entry)
 DT Green fluorescent protein-NFAT1 fusion product.
 XX DE Human, NFAT1 gene; fusion protein; green fluorescent protein; GFP;
 XX KW intracellular signalling; chimera.
 XX OS Aequorea victoria.
 XX OS Homo sapiens.
 XX OS Chimeric.
 XX PN WO9845704-A2.
 XX PD 15-OCT-1998.
 XX PF 07-APR-1998; 98WO-DK00145.
 XX PR 07-APR-1997; 97DK-00000392.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Thastrop O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 XX DR WPI; 1998-594491/50.
 XX DR N-PSDB; AAVT1037.
 XX PT Determining effect on signalling pathways in live cells from
 XX PT redistribution of luminoephores - specifically fusions of green
 XX PT fluorescent protein with a signalling component, and new apparatus,
 XX PT particularly for identifying toxins and potential therapeutic agents.
 XX PS Example 19; Page 258-260; 326p; English.
 XX CC The present sequence represents a green fluorescent protein (GFP)-human
 CC NFAT1 fusion protein. The fusion protein is used in an assay to exemplify
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC luminoephore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (updated on 17-OCT-2003 to standardise OS field)
 CC XX
 XX SQ Sequence 1171 AA;
 XX Query Match 85.4%; Score 1280.5; DB 2; Length 1171;
 XX Best Local Similarity 93.9%; Pred. No. 2.5e-122;
 XX Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;
 QY 1 MWSKGEELFTGVVPIIVELDGVNKHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 Db 1 MWSKGEELFTGVVPIIVELDGVNKHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 QY 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDYL 120
 Db 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDYL 120
 QY 121 VNRTELKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 Db 121 VNRTELKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240

QY 241 LSHGF-----PPEVEEDDG 255
 DB 241 ---GLRSRAMNABERQPPDG 258

RESULT 8
 ABP72573
 ID ABP72573 standard; protein; 357 AA.
 XX ABP72573;
 XX
 DT 23-OCT-2003 (revised)
 DT 29-MAY-2003 (first entry)
 XX
 DE GFP-Drosophila Melted protein pleckstrin homology domain fusion.
 XX
 KM Melted; dmelted; pleckstrin: phosphatidylinositol 5-phosphate; PI(5)P;
 XX binding protein; green fluorescent protein; GFP.
 XX
 OS Aequorea victoria.
 OS Drosophila melanogaster.
 OS Chimeric.

Key Location/Qualifiers
 FT Region 1..245
 FT /label= GFP
 FT 253..350
 FT /label= PH
 FT /note= "dmelted pleckstrin homology domain"
 XX
 XX
 PN MO2003014152-A2.
 XX
 PD 20-FEB-2003.
 XX
 PF 08-AUG-2002; 2002MO-IB003645.
 XX
 PR 09-AUG-2001; 2001GB-00019446.
 XX
 PA (EURO-) EURO MOLECULAR BIOLOGY LAB.
 XX
 PI Cohen S, Teleman A;
 XX
 XX WPI; 2003-256540/25.
 DR
 XX
 XX
 PT New proteins that bind phosphatidylinositol 5-phosphate (PI(5)P), useful
 PT for treating a disease or disorder in which abnormal levels of PI(5)P are
 PT implicated, or visualizing or measuring PI(5)P in vivo or in vitro.
 XX
 XX
 PS Example; Page 22; 41pp; English.

The present sequence is the protein sequence of a fusion protein between
 CC green fluorescent protein (GFP) and the pleckstrin homology (PH) domain
 CC (see also ABP72573) of the Melted protein (dmelted) of Drosophila
 CC melanogaster. Also included in the sequence are short peptide linkers at
 CC the N- and C-termini of the PH moiety. The fusion protein was constructed
 CC and expressed in Drosophila S2 cells in an example from the invention.
 CC The invention relates to proteins that bind to phosphatidylinositol 5-
 CC phosphate (PI(5)P). The PI(5)P binding protein preferably comprises a PH
 CC domain of a Melted protein derived from a vertebrate, a metazoan or an
 CC insect. The binding protein, or a fusion protein comprising it, is useful
 CC for treating a disease or disorder in which abnormal levels of PI(5)P are
 CC implicated, such as diseases that result from dysfunctional signal
 CC transduction and which affect gene expression and cellular proliferation
 CC and differentiation. The PI(5)P binding protein is also useful for
 CC purifying a PI(5)P-containing membrane fraction, or visualizing the
 CC subcellular location of PI(5)P in a cell. Fluorescence microscopy of
 CC transfected Drosophila cells revealed a cortical localisation of the GFP-
 CC dmelted PH fusion construct, showing that PI(5)P is located cortically in
 CC cells. (Updated on 23-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 357 AA;
 Query Match 85.3%; Score 1279.5; DB 6; Length 357;

Best Local Similarity 94.6%; Pred. No. 5.4e-123;
 Matches 245; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 1 MWSKGEELFTGVPLIVELDGVNGHKPSVSGEGDATTGKLTGKICTTGKLPVMPPT 60
 DB 1 MWSKGEELFTGVPLIVELDGVNGHKPSVSGEGDATTGKLTGKICTTGKLPVMPPT 60

QY 61 LVTLITVGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTIFFKDDGNNYTRAEVKEGDTL 120
 DB 61 LVTLITVGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTIFFKDDGNNYTRAEVKEGDTL 120

QY 121 VNRILKGIIDPEKEDNIIIGHLELYNNSHNTYIADKKNIGIKNFKIRHNIEDGSVOLA 180
 DB 121 VNRILKGIIDPEKEDNIIIGHLELYNNSHNTYIADKKNIGIKNFKIRHNIEDGSVOLA 180

QY 181 DHYQONTPIGDGPVLLPNNHYISTQSALSKDPNEKEDMVLLEFYTAGITLGMDELTKK 240
 DB 181 DHYQONTPIGDGPVLLPNNHYISTQSALSKDPNEKEDMVLLEFYTAGITLGMDELTKK 240

QY 241 --LSHGF-----PPEVEEDG 252
 DB 241 GLRSRGLNGSNQPMIEGQ 259

RESULT 9
 AAW85032
 ID AAW85032 standard; protein; 997 AA.
 XX
 XX
 AC AAW85032;
 XX
 DT 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)
 XX
 XX
 DE Green fluorescent protein-IkappaB kinase alpha subunit fusion product.
 XX
 XX
 KM Human; IkappaB kinase gene; alpha subunit; fusion protein;
 KM green fluorescent protein; GFP; intracellular signalling; chimera.
 XX
 OS Aequorea victoria.
 OS Homo sapiens.
 OS Chimeric.

PN MO9845704-A2.
 XX
 PD 15-OCT-1998.
 PD
 XX
 PF 07-APR-1998; 98MO-DK000145.
 PF
 XX
 PR 07-APR-1997; 97DK-00000392.
 PR
 XX
 PA (NOVO) NOVO-NORDISK AS.
 PA
 PI Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 PI N-PSDB; AAV71077.
 DR
 XX
 XX WPI; 1998-594491/50.
 DR
 XX
 XX
 PT Determining effect on signalling pathways in live cells from
 PT redistribution of lumino-phores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.
 XX
 XX
 PS Example 11; Page 231-233; 326pp; English.

The present sequence represents a green fluorescent protein (GFP)-human
 CC IkappaB kinase alpha subunit fusion protein. The fusion protein is used
 CC in an assay to exemplify the invention. The specification describes how
 CC quantitative information about the influence of a molecule on a cellular
 CC response is obtained by recording the variation, caused by the molecule,
 CC on mechanically intact living cells, in the spatially distributed light
 CC emitted from a lumino-phore present in the cells. The variation in light
 CC emission is processed to provide information that correlates spatial
 CC distribution to the degree of the molecule. The method is used to

CC identify agents that (in)directly affect intracellular signalling, especially to screen for potential therapeutic agents or toxins, and to identify new drug targets. (Updated on 17-OCT-2003 to standardise OS field)

CC field)

CC Sequence 997 AA;

Query Match 85.3%; Score 1279; DB 2; Length 997;

Best Local Similarity 90.4%; Pred. No. 2.9e-122; Indels 14; Gaps 1;

Matches 245; Conservative 2; Mismatches 10; Indels 14; Gaps 1;

CC 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60

DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60

QY 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPSGYOERTIFFKDDGNTYKTRAEVFRBPTL 120

DB 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPSGYOERTIFFKDDGNTYKTRAEVFRBPTL 120

QY 121 VNRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180

DB 121 VNRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180

QY 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLLFVTAAGITLGMDELTK 240

DB 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLLFVTAAGITLGMDELTK 240

QY 241 LSHGFPPEVEBDDGTLPMSCAQESGMDRHP 271

DB 240 -----SGLRSAQASNSTMERPP 257

RESULT 10

ID AAW85010 standard; protein; 607 AA.

AC AAW85010;

DT 17-OCT-2003 (revised)

DT 08-FEB-1999 (first entry)

DE p38-green fluorescent protein fusion product.

KW Human; p38 gene; fusion protein; green fluorescent protein; GFP;

KM intracellular signalling; chimera.

OS Aequorea victoria.

OS Homo sapiens.

OS Chimeric.

XX W09845704-A2.

PD 15-OCT-1998.

PF 07-APR-1998; 98WO-DK000145.

PR 07-APR-1997; 97DK-00000392.

PA (NOVO) NOVO-NORDISK AS.

PI Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;

DR WPI; 1998-594491/50.

DR N-PSDB; AAV71025.

XX Determining effect on signalling pathways in live cells from

PT redistribution of lymphocytes - specifically fusions of green

PT fluorescent protein with a signalling component, and new apparatus,

PT particularly for identifying toxins and potential therapeutic agents.

XX Example 8; Page 90-91; 326pp; English.

XX The present sequence represents a human p38-green fluorescent protein

CC fusion product. The fusion protein is used in an assay that exemplifies the invention. The specification describes how quantitative information about the influence of a molecule on a cellular response is obtained by recording the variation, caused by the molecule, on mechanically intact living cells, in the spatially distributed light emitted from a lymphocyte present in the cells. The variation in light distribution is processed to provide information that correlates spatial distribution to the degree of the molecule. The method is used to identify agents that (in)directly affect intracellular signalling, especially to screen for potential therapeutic agents or toxins, and to identify new drug targets. (Updated on 17-OCT-2003 to standardise OS field)

CC Sequence 607 AA;

Query Match 85.2%; Score 1278; DB 2; Length 607;

Best Local Similarity 84.7%; Pred. No. 1.7e-122; Indels 24; Gaps 3;

Matches 249; Conservative 6; Mismatches 15; Indels 24; Gaps 3;

CC 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60

DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60

QY 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPSGYOERTIFFKDDGNTYKTRAEVFRBPTL 120

DB 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPSGYOERTIFFKDDGNTYKTRAEVFRBPTL 120

QY 121 VNRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180

DB 121 VNRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180

QY 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLLFVTAAGITLGMDELTK 239

DB 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLLFVTAAGITLGMDELTK 240

QY 240 -----KLSHGFP-----EVEBDDGTLPMSCAQESGMDRHPAACS 276

DB 241 GLRSKMSQCPPTFYRQELNKTITWEVERYONLSPV-----SGAGSYCAA 288

RESULT 11

ID AAY85574 standard; protein; 2608 AA.

AC AAY85574;

DT 07-JUN-2000 (first entry)

DE Hs-UNC-53/3/GFP fusion insert of plasmid pG13303.

KW UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;

KM anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;

KW antisclerotic; antimetastatic; anti-arthritis; autoimmune disease.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1194 /note= "unspecified"

PN W09963080-A1.

PD 09-DEC-1999.

PF 02-JUN-1999; 99WO-EP003848.

PR 03-JUN-1998; 98GB-00011962.

PA (JANC) JANSSEN PHARM NV.

PI Luyten WHM, De Raeymaeker MC, Geysen JGH, Bogaert THOE;

PI Maerten LJS, Verhaesele P, Van De Craen M;

XX

DR WPI; 2000-116370/10.
 DR N-PSDB; AAA07846.
 XX
 PT Novel proteins and nucleic acids e.g. for treating neurodegeneration.
 XX
 PS Disclosure; Fig 7e; 146pp; English.
 XX
 CC The invention provides vertebrate (human) protein homologue of a UNC-53
 CC protein of *Caenorhabditis elegans*. The UNC-53 binds to microtubules or
 CC their plus ends. The UNC-53 sequences are used to promote neural
 CC regeneration, revascularization and wound healing; also for treating
 CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
 CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
 CC polynucleotides can be used for recombinant production of the proteins,
 CC as a source of probes for detecting allelic variants and polymorphisms,
 CC for sequencing genomic DNA and for detecting UNC-53 expression; and as
 CC source of therapeutic antisense sequences. Cells that express the protein
 CC are used to identify regulators of cell shape, growth, motility and
 CC migration. They can also be used to identify proteins that are involved
 CC in signal transduction pathways also involving UNC-53, and to identify
 CC compounds that alter attachment of UNC-53 to microtubules. A target gene
 CC coupled to a UNC-53 encoding sequence may be used to deliver the target
 CC gene to a cellular microtubule or its plus ends. The present sequence
 CC represents the amino acid sequence of a full-length Hs-UNC-53/3 in fusion
 CC with GFP insert of plasmid pG13305
 CC
 XX
 SQ Sequence 2608 AA;
 Query Match 85.2%; Score 1278; DB 3; Length 2608;
 Best Local Similarity 91.4%; Pred. No. 1.5e-121; Mismatches 5; Indels 6; Gaps 2;
 Matches 245; Conservative 5; Mismatches 12; Indels 6; Gaps 2;
 QY 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYGLTLKFICTTGKLPVPMPT 60
 DB 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYGLTLKFICTTGKLPVPMPT 60
 QY 61 LVTTLTLYGQCFRRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNYKTRAVYKRGDTL 120
 DB 61 LVTTLTLYGQCFRRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNYKTRAVYKRGDTL 120
 QY 121 VNRLEKGIIDFKEDGNILGHKLEYNVNSHNYITMADKQNGIKVNFKIRNIEDGSVOLA 180
 DB 121 VNRLEKGIIDFKEDGNILGHKLEYNVNSHNYITMADKQNGIKVNFKIRNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLLFEYTAAGITLGMDELYK 239
 DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLLFEYTAAGITLGMDELYK 240
 QY 240 -KLSH---GFPEVEEODDGTIPMSGCA 262
 DB 241 SDLEHMPVLGVASKRQPAVGSKRVHTA 268
 DB 241 SDLEHMPVLGVASKRQPAVGSKRVHTA 268
 RESULT 12
 ADB90657
 ID ADB90657 standard; protein; 1089 AA.
 XX
 AC ADB90657;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human and rat chimeric protein #2.
 XX
 XX nuclear translocation protein; glucocorticoid receptor DNA; DNA chimera;
 KW Cytostatic; Cardiac; Antiartherosclerotic; Antiarthritic;
 KM Antiinflammatory; Gene therapy; cancer; heart disease; arteriosclerosis;
 KM arthritis; inflammatory states; human; rat.
 XX
 XX Chimeric.
 OS Homo sapiens.
 OS Rattus sp.
 XX
 PN MO2003027639-A2.

XX
 PD 03-APR-2003.
 XX
 PF 30-SEP-2002; 2002MO-US031043.
 XX
 PR 28-SEP-2001; 2001US-0325178P.
 XX
 XX (HAGER/) HAGER G L.
 XX
 PI Hager GL, Mackem S;
 XX
 DR WPI; 2003-402984/38.
 XX
 PT Making a recombinant nuclear translocation protein, useful for treating
 PT cancer, heart disease, by covalently connecting glucocorticoid receptor
 PT DNA, superfamily receptor DNA and nucleic acid sequence for a marker
 PT protein domain.
 PS
 PS Disclosure; Page 75-81; 115pp; English.
 XX
 CC The present sequence relates to making a recombinant nuclear
 CC translocation protein by covalently connecting a glucocorticoid receptor
 CC DNA sequence coding for the cytoplasmic/nuclear translocation domain of
 CC the glucocorticoid receptor protein, a superfamily receptor DNA sequence
 CC coding for the ligand binding domain of the superfamily receptor protein,
 CC and a nucleic acid sequence for a marker protein domain, forming a DNA
 CC chimera. The methods, chimera, and protein are useful for treating a
 CC defective translocation of a superfamily receptor protein from the
 CC cytoplasm to the nucleus of the cell in an animal, such as cancer, heart
 CC disease, arteriosclerosis, arthritis, and inflammatory states. The
 CC present sequence represents a rat/human chimera of the invention.
 CC
 XX
 SQ Sequence 1089 AA;
 Query Match 85.1%; Score 1277; DB 7; Length 1089;
 Best Local Similarity 86.4%; Pred. No. 5.2e-122; Mismatches 25; Indels 6; Gaps 2;
 Matches 247; Conservative 8; Mismatches 25; Indels 6; Gaps 2;
 QY 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYGLTLKFICTTGKLPVPMPT 60
 DB 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYGLTLKFICTTGKLPVPMPT 60
 QY 61 LVTTLTLYGQCFRRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNYKTRAVYKRGDTL 120
 DB 61 LVTTLTLYGQCFRRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNYKTRAVYKRGDTL 120
 QY 121 VNRLEKGIIDFKEDGNILGHKLEYNVNSHNYITMADKQNGIKVNFKIRNIEDGSVOLA 180
 DB 121 VNRLEKGIIDFKEDGNILGHKLEYNVNSHNYITMADKQNGIKVNFKIRNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLLFEYTAAGITLGMDELYK 240
 DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLLFEYTAAGITLGMDELYK 240
 QY 241 -----LSHGFPPEVEEODDGTIPMSGCAQSGSDRHPACASARIV 281
 DB 241 GALDQSKESLAPPGRDEVGSL-LGQGRGVYDFFKSLRGATVAV 285
 DB 241 GALDQSKESLAPPGRDEVGSL-LGQGRGVYDFFKSLRGATVAV 285
 RESULT 13
 ADE24110
 ID ADE24110 standard; protein; 284 AA.
 XX
 AC ADE24110;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Discosoma sp DsRed1 reporter protein #2.
 XX
 XX site-specific DNA recombination; Cre recombinase;
 KW protein-transduction domain; gene activation; gene inactivation;
 KW chromosomal translocation; reporter protein.
 XX

OS	Synthetic.
OS	Discosoma sp.
PN	WO2003070931-A2.
XX	
PD	28-AUG-2003.
XX	
PF	19-FEB-2003; 2003WO-EP001680.
XX	
PR	21-FEB-2002; 2002DE-01007313.
FR	16-JUL-2002; 2002DE-01032196.
XX	
PA	(VISI-) VISION 7 GMBH.
XX	
PI	Baum C, Wall E, Osterlag W, Klump H, Schiedlmeier B;
DR	WPT; 2003-767353/72.
XX	
PT	In vitro or in vivo site-specific DNA recombination, useful e.g. for gene
PR	inactivation, using Cre recombinase that lacks heterologous protein-
XX	transduction domain.
PS	Claim 22; SEQ ID NO 7; 84pp; German.
XX	
CC	This invention describes a novel method of site-specific DNA
CC	recombination in eukaryotic cells in vitro comprising using a
CC	bacteriophage P1 Cre recombinase that lacks heterologous protein-
CC	transduction domains. The invention also contains a reporter system for
CC	detecting site-specific DNA recombination in eukaryotic cells.
CC	Recombination with Cre recombinase lacking heterologous protein-
CC	transduction domains is used for therapeutic site-specific recombination
CC	in eukaryotic cells, in vivo or in vitro, e.g. for partial inactivation
CC	or activation of genes, also to create a chromosomal translocation. The
CC	modified recombinase can enter intact (not electrically or chemically
CC	treated) cells when added to culture medium and catalyzes recombination
CC	events in the nucleus with recombination rate over 50% after a single
CC	application. The Cre-recombinase does not need to be supplied from
CC	nucleic acid introduced into the cell, so its effect is immediate but
CC	short-lived, limiting cytotoxic effects; eliminating the risk of non-
CC	-specific integration of DNA into the genome and providing quick results.
CC	This sequence represents a Discosoma sp reporter protein DsRed1 which
CC	contains a fragment of green fluorescent protein (eGFP) used in the
CC	method described in the invention.
XX	
SQ	Sequence 284 AA:
Query Match	85.1%; Score 1276.5; DB 7; Length 284;
Best Local Similarity	94.2%; Pred. No. 7.8e-123;
Matches 243; Conservative	2; Mismatches 10; Indels 3; Gaps 1
QY	2 VSKGEELFTGVPLIIVELDGDVNGHKFSVSGEGDATYGLTKLKEICTTGKLPVPWPMTL 61
DB	18 VSKGEELFTGVPLIIVELDGDVNGHKFSVSGEGDATYGLTKLKEICTTGKLPVPWPMTL 77
QY	62 VTTLTVYGOCFERYDHNKKOHDPFSAMEGVQORRTFFPKDGNYKTAAYKFKEDTLY 121
DB	78 VTTLTVYGOCFERYDHNKKOHDPFSAMEGVQERTTFPKDGNYKTAAYKFKEDTLY 137
QY	122 NRIELKGIIDPEKDGNILGHKLLEYNNSHNVYIMADOKNGIKIVNFKIRINIEDGSVOLAD 181
DB	138 NRIELKGIIDPEKDGNILGHKLLEYNNSHNVYIMADOKNGIKIVNFKIRINIEDGSVOLAD 197
QY	182 HYQQNTPIIGDGGVLLPDDHYILLSTQSALSDPNEKRDHNVILLEFYVAAGITTLGMDELYK-- 239
DB	198 HYQQNTPIIGDGGVLLPDDHYILLSTQSALSDPNEKRDHNVILLEFYVAAGITTLGMDELYXSG 257
QY	240 -KLSHGFPEVEEODDGT 256
DB	258 LRSSADNPASTTNKDAQT 275

ID	ADB90661 standard; protein; 1090 AA.
XX	
AC	ADB90661;
DT	04-DEC-2003 (first entry)
XX	
DE	Human and rat chimeric protein #4.
XX	
KW	nuclear translocation protein; glucocorticoid receptor DNA; DNA chimera;
KM	Cyclostatic Cardiant; Antiarteriosclerotic; Arthritis;
KV	Antiinflammatory; Gene therapy; cancer; heart disease; arteriosclerosis;
XX	arthritis; inflammatory states; human; rat.
OS	Chimeric.
OS	Homo sapiens.
XX	Rattus sp.
PN	MO2003027639-A2.
PD	03-APR-2003.
XX	
PF	30-SEP-2002; 2002WO-US031043.
PR	28-SEP-2001; 2001US-0325178P.
XX	
PA	(HAGE/) HAGER G L.
PI	Hager GL; Mackem S;
XX	WPI; 2003-402984/38.
DR	N-PSDB; ADB90660.
XX	
PT	Making a recombinant nuclear translocation protein, useful for treating
PT	cancer, heart disease, by covalently connecting glucocorticoid receptor
PT	DNA, superfamily receptor DNA and nucleic acid sequence for a marker
PT	protein domain.
XX	
P5	Disclosure; Page 108-114; 115pp; English.
CC	
CC	The present sequence relates to making a recombinant nuclear
CC	translocation protein by covalently connecting a glucocorticoid receptor
CC	DNA sequence coding for the cytoplasmic/nuclear translocation domain of
CC	the glucocorticoid receptor protein, a superfamily receptor DNA sequence
CC	coding for the ligand binding domain of the superfamily receptor protein,
CC	and a nucleic acid sequence for a marker protein domain, forming a DNA
CC	chimera. The methods, chimera, and protein are useful for treating
CC	defective translocation of a superfamily receptor protein from the
CC	cytoplasm to the nucleus of the cell in an animal, such as cancer, heart
CC	disease, arteriosclerosis, arthritis, and inflammatory states. The
CC	present sequence represents a rat/human chimera of the invention.
XX	
SQ	Sequence 1090 AA:
Query Match	85.1%; Score 1276.5; DB 7; Length 1090;
Best Local Similarity	86.1%; Pred. No. 5.9e-122;
Matches 247; Conservative	8; Mismatches 25; Indels 7; Gaps 2.
QY	1 MWSKGEELFTGVNVLIVELDQVNGHFRFSVSGEGEDPATYGLTLKFICTTGTLPVPWP 60
DB	1 MWSGDELFLGVNVLIVELDQDVNGHFRFSVSGEGEDPATYGLTLKFICTTGTLPVPWP 60
QY	61 LVTLTYTGVCQFSRPDHMQOHDFEKKAMPBGVYOERTIFFKDDGNTRYAEVFPESDYL 120
DB	61 LVTLTYTGVCQFSRPDHMQOHDFEKKAMEGGVYOERTIFFKDDGNTRYAEVFPESDYL 120
QY	121 VNRITELGIPEKEKGNTLGHLEENYNASHNYIMADOKNGIKYNPKIRHIIEGSVOLA 180
DB	121 VNRITELGIPEKEKGNTLGHLEENYNASHNYIMADOKNGIKYNPKIRHIIEGSVOLA 180
QY	181 DHYOQNTPIGDGPVLLPDNHVLTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB	181 DHYQONTPIGDGPVLLPDNHVLTQSALSKDPNEKRDMVLLFEVTAAAGITLGMDELYKS 240

QY 241 -----LSHGPPPEVEODDGLPMSCAOSGMDRHPAACASARINV 281
 Db 241 GALLIDSKESLIAPGRDEYVGSIL-IGQHGSGVMDPFKSLRGKATYAV 286

RESULT 15

ADE28562
 ID ADE28562 standard; protein; 286 AA.

AC ADE28562;

DT 29-JAN-2004 (first entry)

DE EGFP/ human CD1a fusion construct protein.

XX CDI fusion protein; endosomal targeting; antibacterial; virucide;
 XX cytostatic; antidiabetic; antiinflammatory; dermatological;
 KW immunosuppressive; neuroprotective; immune response; microbial;
 KW viral infection; autoimmune disorder; diabetes; lupus;
 KW multiple sclerosis; cancer; gene therapy; vaccine; human; EGFP/CD1a;
 KW chimeric; enhanced green fluorescent protein.

OS Unidentified.

OS Chimeric.

OS Homo sapiens.

PN WO2003066820-A2.

XX 14-AUG-2003.

PD 05-FEB-2003; 2003WO-US003550.

PF 05-FEB-2002; 2002US-0355432P.

PR (REGC) UNIV CALIFORNIA.

PA Modlin RL;

PI WPI; 2003-748121/70.

DR N-PSDB; ADE28561.

XX New chimeric nucleic acid molecule encoding a CDI fusion protein, useful
 PT for inducing an immune response against antigens associated with, e.g.
 PT microbial or viral infections, autoimmune disorders or cancer.

PS Claim 26; Fig 11A; 114pp; English.

XX The invention relates to a novel chimeric nucleic acid molecule
 CC consisting of a nucleotide sequence encoding a CDI fusion protein
 CC comprising a CDI endosomal targeting sequence or its fragment and an
 CC antigen of interest. The polynucleotide of the invention demonstrates
 CC antibacterial, virucide, cytostatic, antidiabetic, antiinflammatory,
 CC dermatological, immunosuppressive and neuroprotective activities whilst
 CC the composition and methods may be useful for inducing an immune response
 CC against antigens of interest which are associated with microbial or viral
 CC infections, autoimmune disorders including diabetes, lupus and multiple
 CC sclerosis and antigens associated with cancers. Furthermore, the
 CC composition may be utilised in gene therapy and vaccine development. The
 CC current sequence is that of the EGFP (enhanced green fluorescent
 CC protein)/ human CD1a fusion construct protein of the invention.

XX Sequence 286 AA;

Query Match 85.1%; Score 1276; DB 7; Length 286;

Best Local Similarity 99.2%; Pred. No. 8,8e-123;

Matches 239; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWSKGEELFTGVVPLIVELDGVNNGHRSVSGEGSDATYGLTLKFTCTGKLPVWPPT 60

Db 20 MWSKGEELFTGVVPLIVELDGVNNGHRSVSGEGSDATYGLTLKFTCTGKLPVWPPT 79

QY 61 LVTLLTYGVQCFSRYPDMKQHDFFKSAHPBGVYQERTIFFKDDGNKYKTRAEVKEEGDTL 120

Db 80 LVTLLTYGVQCFSRYPDMKQHDFFKSAHPBGVYQERTIFFKDDGNKYKTRAEVKEEGDTL 139
 QY 121 VNRILELKGIDFKEDGNITIGHKLEIYNVNSHNYIIMADKQKGIKXNFKIRHNIEDGSVOLA 180
 Db 140 VNRILELKGIDFKEDGNITIGHKLEIYNVNSHNYIIMADKQKGIKXNFKIRHNIEDGSVOLA 199
 QY 181 DRYQONTPIGDGPVILLPNHYLSTQSALSCKDPNEKRDHMYLLEFVTAGITLGMDELTKK 240
 Db 200 DRYQONTPIGDGPVILLPNHYLSTQSALSCKDPNEKRDHMYLLEFVTAGITLGMDELTKK 259
 QY 241 L 241
 Db 260 I 260

Search completed: May 14, 2004, 08:25:16
 Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 / Search time 21 Seconds
(without alignments)

1287,134 Million cell updates/sec

Title: US-09-931-232-1

Sequence: 1 MVSKEBELFTGVVPIIVLELD.....AQESGMDRHPACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

PIR 78:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124.5	83.0	238	1 JQ1514	green-fluorescent
2	221.5	14.8	461	1 DCM50	ornithine decarbox
3	221.5	14.8	461	2 I56477	ornithine decarbox
4	210.5	14.0	461	2 I55356	ornithine decarbox
5	208.5	13.9	461	1 DCRTO	ornithine decarbox
6	200	13.3	455	1 DCHYOC	ornithine decarbox
7	154.5	10.3	461	1 DCHHO	ornithine decarbox
8	111.5	7.4	450	1 DCHHO	ornithine decarbox
9	103	6.9	460	2 A43563	ornithine decarbox
10	101	6.7	785	2 H72228	ornithine decarbox
11	99.5	6.6	1224	1 ERUDAH	coatomer complex a
12	95.5	6.4	4307	2 T20721	hypothetical prote
13	92.5	6.2	752	1 KXRTCI	protoprotein convert
14	92	6.1	336	2 C64468	hypothetical prote
15	91	6.1	357	2 G81355	tRNA (tyracil-5-)-m
16	91	6.0	632	2 T06586	DNA-binding protei
17	90.5	6.0	861	2 H64102	leucine-tRNA 119as
18	90.5	6.0	1193	2 F68698	DNA-directed RNA p
19	90	6.0	1039	2 H95115	conserved hypothet
20	89	5.9	433	2 B90103	synaptotagmin c-p65
21	88.5	5.9	502	2 JH0414	T-complex protein
22	88.5	5.9	578	1 I40794	dihydrolycanide d
23	88.5	5.9	887	2 B82590	leucyl-tRNA synthe
24	88.5	5.9	889	2 JCS576	inner-alpha-trypsi
25	88.5	5.9	6642	2 T27577	protein UNC-89 - C
26	88	5.9	471	2 T27856	hypothetical prote
27	88	5.9	1259	2 AE1055	probable exported
28	87.5	5.8	2573	2 D71614	hypothetical prote
29	87	5.8	281	2 AD2052	hypothetical prote

30	86.5	5.8	370	2 E70390	iron-sulfur cofact
31	86.5	5.8	441	2 T06936	photosystem II chl
32	86.5	5.8	860	2 AC0582	leucyl-tRNA synthe
33	86.5	5.8	874	2 JC4930	S-layer protein pr
34	86.5	5.8	2222	1 A36028	DNA-directed DNA p
35	86	5.7	263	2 S53488	water-stress-induc
36	86	5.7	1039	2 D97985	hypothetical prote
37	86	5.7	1290	2 S76853	hypothetical prote
38	85.5	5.7	1280	2 D83917	DNA topoisomerase
39	85.5	5.7	788	1 JDVLDH	DNA-directed DNA p
40	85.5	5.7	877	2 AC2211	home transport pro
41	85.5	5.7	1134	2 A60234	IGA Fc receptor pr
42	85.5	5.7	1164	1 FCS0AG	IGA Fc receptor pr
43	85	5.7	331	2 T41758	AcAMPV orf13 - Bom
44	85	5.7	466	2 B64207	heterocyst maturat
45	85	5.7	3744	2 S46715	hypothetical prote

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C/Species: Aequorea victoria

C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #ext change 23-Mar-2001

C/Accession: US0692; JQ1514; P00335; S48693; S51330; S51331

R/Prasner, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.

Gene 111, 229-233, 1992

A/Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A/Reference number: JQ1514; MUID:192175527; PMID:1347277

A/Accession: J50692

A/Molecule type: DNA

A/Residues: 1-107, 'S', 109-238 <PRA1>

A/Cross-references: GB:M62654; NID:G155662; PIDN:AAA27722.1; PID:G155663

A/Accession: JQ1514

A/Molecule type: mRNA

A/Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>

A/Cross-references: GB:M62653; NID:G155660; PIDN:AAA27721.1; PID:G155661

A/Accession: P00335

A/Molecule type: Protein

A/Residues: 46-64/74-122/132-151/154-183/185-200 <PRA3>

R/Inouye, S.; Teuil, F.I.

FEBS Lett. 351, 211-214, 1994

A/Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A/Reference number: S48693; MUID:94364470; PMID:8082767

A/Accession: S48693

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>

R/Matkins, J.N.; Campbell, A.K.

submitted to the EMBL Data Library, January 1995

A/Reference number: S51330

A/Accession: S51330

A/Molecule type: mRNA

A/Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 229-238 <INO>

A/Cross-references: EMBL:X83959; NID:9634008; PIDN:CAA58789.1; PID:9634009

A/Experimental source: clone gfp2

A/Accession: S51331

A/Molecule type: mRNA

A/Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 210-238 <INO>

A/Cross-references: EMBL:X83960; NID:9634010; PIDN:CAA58790.1; PID:9634011

A/Experimental source: clone gfp2

R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A/Reference number: A65692; PDB:1GFL

A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-95

A/Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A/Title: The molecular structure of green fluorescent protein.

A/Reference number: A58953; MUID:96294543; PMID:9631087

A:Contents: annotation; X-ray crystallography, 1.9 angstroms
 C:Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJNV) emitting
 C:Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
 C:Genetics:
 A:Gene: GFP
 A:Inserts: 69/3; 167/3
 C:Superfamily: green-fluorescent protein
 C:Keywords: chromoprotein; luminescence
 F:65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
 F:66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 83.0%; Score 1245; DB 1; Length 238;
 Best Local Similarity 97.1%; Pred. No. 1.5e-93;
 Matches 231; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKSEELFTGVVPLIVELDGVNCHKFSVSGEGSDATYGLTLKFLCTTGKLPVPPPTL 61
 Db 1 MSKEELFTGVVPLIVELDGVNCHKFSVSGEGSDATYGLTLKFLCTTGKLPVPPPTL 60
 QY 62 VTTLTYGVQCFSRYPDMKQHDPEKSAFPGYVQERTTFKCDGNYTRARVKEGDTLV 121
 Db 61 VTTFYSYGVQCFSRYPDMKQHDPEKSAFPGYVQERTTFKCDGNYTRARVKEGDTLV 120
 QY 122 NRLEIKDIDREKGNITLGHLEKLYNNHNYIMADKQNGIKYFKIRNIEDGSVOLAD 181
 Db 121 NRLEIKDIDREKGNITLGHLEKLYNNHNYIMADKQNGIKYFKIRNIEDGSVOLAD 180
 QY 182 HYQONTPIGDGFPVLLPNNHYLSTQSALSKDPEKEDRMVLEFPTAGITLGMDELYK 239
 Db 181 HYQONTPIGDGFPVLLPNNHYLSTQSALSKDPEKEDRMVLEFPTAGITLGMDELYK 238

RESULT 2
 DCM50
 Ornithine decarboxylase (EC 4.1.1.17) [validated] - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 27-Nov-1985 #sequence revision 27-Nov-1985 #text change 18-Aug-2000
 C:Accession: A01077; S02124; A25427; B25427; I55283; S34065
 R:Kahana, C.; Nathans, D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 1673-1677, 1985
 A:Title: Nucleotide sequence of murine ornithine decarboxylase mRNA.
 A:Reference number: A01077; MUID:8516193; PMID:3856848
 A:Accession: A01077

A:Molecule type: mRNA
 A:Residues: 1-461 <RAH>
 A:Cross-references: GB:M10624; NID:g200119; PIDN:AAA39845.1; PID:g200120
 R:Gupta, M.; Coffino, P.
 J. Biol. Chem. 260, 2941-2944, 1985
 A:Title: Mouse ornithine decarboxylase. Complete amino acid sequence deduced from cDNA.
 A:Reference number: A22295; MUID:85131068; PMID:2982844
 A:Accession: A22295

A:Molecule type: mRNA
 A:Residues: 1-461 <GUP>
 A:Cross-references: GB:M19521; GB:M20617; GB:J03615; NID:g200127; PIDN:AAA51638.1; PID:g200127
 R:Coffino, P.; Chen, E. L.
 Nucleic Acids Res. 16, 2731-2732, 1988
 A:Title: Nucleotide sequence of the mouse ornithine decarboxylase gene.
 A:Reference number: S02124; MUID:88203223; PMID:3362685
 A:Accession: S02124

A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-461 <COF>
 A:Cross-references: EMBL:X07392; NID:g53517; PIDN:CAA30301.1; PID:g53518
 R:Richok, N.J.; Seppanen, P.U.; Kontula, K.K.; Janne, P.A.; Bardin, C.W.; Janne, O.A.
 Proc. Natl. Acad. Sci. U.S.A. 83, 594-598, 1986
 A:Title: Two ornithine decarboxylase mRNA species in mouse kidney arise from size hetero
 A:Reference number: A94117; MUID:86120987; PMID:3456155
 A:Accession: A25427

A:Molecule type: mRNA
 A:Residues: 55-56,180-205, 'E', 207-461 <HIC1>
 A:Cross-references: GB:M12330; NID:g200121; PIDN:AAA39846.1; PID:g387499
 A:Note: this mRNA was designated as clone podc16
 A:Accession: B25427

A:Molecule type: mRNA
 A:Residues: 320-349, 'H', 351-461 <HIC2>
 A:Cross-references: GB:M12331; NID:g200130; PIDN:AAA39848.1; PID:g200131
 A:Note: this mRNA was designated as clone podc74
 A:Note: the authors translated the codon CAT for residue 330 as Tyr
 J.Poulin, R.; Lu, L.; Ackermann, B.; Bey, P.; Pegg, A.E.
 J. Biol. Chem. 267, 150-158, 1992
 A:Title: Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by
 A:Reference number: A41744; MUID:92112641; PMID:130582
 A:Contents: annotation; active site

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-461 <RES>
 A:Cross-references: GB:J03733; NID:g200132; PIDN:AAA39849.1; PID:g387501
 R:Teitka, S.R.; Turk, C.W.; Coffino, P.
 Biochem. J. 293, 285-295, 1993
 A:Title: Multiple active conformers of mouse ornithine decarboxylase.
 A:Reference number: S34065; MUID:93319524; PMID:8328969
 A:Accession: S34065

A:Molecule type: protein
 A:Residues: 357-359, 'X', 361-367 <TS1>
 C:Genetics:
 A:Inserts: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
 A:Complex: homodimer
 C:Function:
 A:Description: catalyzes the decarboxylation of ornithine to putrescine
 A:Pathway: polyamine biosynthesis
 A:Note: the first and rate-limiting reaction in the pathway
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; homodimer; phosphoprotein; polyamine bio
 F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental
 F:360/Active site: Cys (shared with dimeric partner) #status experimental

Query Match 14.8%; Score 221.5; DB 1; Length 461;
 Best Local Similarity 87.8%; Pred. No. 3.2e-10;
 Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 234 MDELYKTL-SHGFPEVEHODDGLTPMSCAQSGMDRHPAACAARINV 281
 Db 413 MMQIMKQIQSHGFPEVEHODDGLTPMSCAQSGMDRHPAACAARINV 461

RESULT 3
 156477
 Ornithine decarboxylase (EC 4.1.1.17) - mouse
 C:Species: Mus musculus domesticus (western European house mouse)
 C:Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text change 08-Sep-2000
 C:Accession: I56477
 R:Joannes, G.U.; Berger, F.G.
 J. Mol. Evol. 36, 555-567, 1993
 A:Title: Domains within the mammalian ornithine decarboxylase messenger RNA have evolved
 A:Reference number: I56477; MUID:93353527; PMID:8350350
 A:Accession: I56477

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-461 <JOH>
 A:Cross-references: GB:S64539; NID:g404561; PIDN:AB27809.1; PID:g404562
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis; F
 F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 14.8%; Score 221.5; DB 2; Length 461;
 Best Local Similarity 87.8%; Pred. No. 3.2e-10;
 Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 224 MDELYKTL-SHGFPEVEHODDGLTPMSCAQSGMDRHPAACAARINV 281
 Db 413 MMQIMKQIQSHGFPEVEHODDGLTPMSCAQSGMDRHPAACAARINV 461

RESULT 4
155356
ornithine decarboxylase (EC 4.1.1.17) - shrew mouse
C/Species: Mus pahari
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Sep-2000
C/Accession: 155356
R/Johannes, G.; Berger, F.G.
J. Biol. Chem. 267, 10108-10115, 1992
A/Title: Alterations in mRNA translation as a mechanism for the modification of enzyme
A/Reference number: 155356; MUID:92250571; PMID:1374399
A/Accession: 155356
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-461 <RES>
A/Cross-references: GB:M87223; NID:g200123; PIDN:AAA39847.1; PID:g200124
C/Genetics:
A/Comment: This pyridoxal phosphate enzyme, which belongs to a multigene family, catalyzes
C/Superfamily: ornithine decarboxylase
C/Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
F/69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
Query Match 14.0%; Score 210.5; DB 2; Length 461;
Best Local Similarity 29.1%; Pred. No. 2.5e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGYVQ---ERTFFKDDGNKYTRAIVKEGDTLVNRIELKG-----IDFKEDGN 136
DB 205 PETFYQAVSDARCVF--DWG-----TEVGFSS---MYLIDIGGPGSGEDTKIKFEETTS 253
QY 137 ILGHKLE-YNNVSNVNYTMADKQNGIKVNFKIRNI-----EDGSYQ 179
DB 254 VINPALDKYFPDSGVRILIAEGRYVVASAFTLANIILAKTWKEQGSDEDESNQET 313
QY 180 ADHYQNTPTIGDGPVLLPDNHYLSTQSALSKDPNEK-----RDHMY----- 220
DB 314 EMYYNDGVYGSFNCILVDHAHVKALLQKRPDPDEKYSSISWGPTCDGLDRIYERCSLP 373
QY 221 -----LLE---FVTAAGITL-----MDLYKKL-SHGFPPEVEROD 253
DB 374 EMHVGDDMWLFENMGAYTVAASTFNGFQRPNTIYVMSRMWQLMKRIQSHGFPPEVEROD 433
QY 254 DGTLPMSCAQESGMDRHPAACASARINV 281
DB 434 DGTLPMSCAQESGMDHSAACASARINV 461
RESULT 5
DCHTCC
ornithine decarboxylase (EC 4.1.1.17) - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 18-Jun-1999
C/Accession: A27361; A33710; C33710; S01406; S14352
R/van Kranen, H.J.; van de Zande, L.; van Kreijl, C.F.; Bisschop, A.; Wieringa, B.
Gene 60, 145-155, 1987
A/Title: Cloning and nucleotide sequence of rat ornithine decarboxylase cDNA.
A/Reference number: A27361; MUID:88167817; PMID:3443298
A/Accession: A27361
A/Molecule type: mRNA
A/Residues: 1-461 <VAN>
A/Cross-references: GB:M16982; NID:g205803; PIDN:AAA41737.1; PID:g205804; GB:M19157
A/Experimental source: Clone pODC.F10
R/Wen, L.; Huang, J.K.; Blackshear, P.J.
J. Biol. Chem. 264, 9016-9021, 1989
A/Title: Rat ornithine decarboxylase gene. Nucleotide sequence, potential regulatory ele
A/Reference number: A33710; MUID:89255376; PMID:2722815
A/Accession: A33710
A/Molecule type: DNA
A/Residues: 1-461 <MEN1>
A/Cross-references: GB:J04792
A/Accession: C33710
A/Molecule type: mRNA

A/Residues: 1-461 <MEN2>
A/Cross-references: GB:J04791; NID:g205807; PIDN:AAA6164.1; PID:g205808
R/van Steeg, H.; van Oostrom, C.T.M.; van Kranen, H.J.; van Kreijl, C.F.
Nucleic Acids Res. 16, 8173-8174, 1988
A/Title: Nucleotide sequence of the rat ornithine decarboxylase gene.
A/Reference number: S01406; MUID:88335556; PMID:3419906
A/Accession: S01406
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-461 <VAN2>
A/Cross-references: EMBL:X07944; NID:g56786; PIDN:CAA30765.1; PID:g56787
R/van Steeg, H.; van Oostrom, C.T.M.; Hodemaekers, H.M.; Peters, L.; Thomas, A.A.M.
Biochem. J. 274, 521-526, 1991
A/Title: The translation in vitro of rat ornithine decarboxylase mRNA is blocked by its
A/Reference number: S14352; MUID:91174765; PMID:2006916
A/Accession: S14352
A/Molecule type: mRNA
A/Residues: 1-37 <STE>
A/Comment: This pyridoxal phosphate enzyme, which belongs to a multigene family, catalyzes
C/Genetics:
A/Insertions: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
C/Superfamily: ornithine decarboxylase
C/Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
F/69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
F/360/Active site: Cys (shared with dimeric partner) #status predicted
Query Match 13.9%; Score 208.5; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 3.6e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGYVQ---ERTFFKDDGNKYTRAIVKEGDTLVNRIELKG-----IDFKEDGN 136
DB 205 PETFYQAVSDARCVF--DWG-----TEVGFSS---MYLIDIGGPGSGEDTKIKFEETTS 253
QY 137 ILGHKLE-YNNVSNVNYTMADKQNGIKVNFKIRNI-----EDGSYQ 179
DB 254 VINPALDKYFPDSGVRILIAEGRYVVASAFTLANIILAKTWKEQGSDEDESNQET 313
QY 180 ADHYQNTPTIGDGPVLLPDNHYLSTQSALSKDPNEK-----RDHMY----- 220
DB 314 EMYYNDGVYGSFNCILVDHAHVKALLQKRPDPDEKYSSISWGPTCDGLDRIYERCSLP 373
QY 221 -----LLE---FVTAAGITL-----MDLYKKL-SHGFPPEVEROD 253
DB 374 EMHVGDDMWLFENMGAYTVAASTFNGFQRPNTIYVMSRMWQLMKRIQSHGFPPEVEROD 433
QY 254 DGTLPMSCAQESGMDRHPAACASARINV 281
DB 434 VGTLPMSCAQESGMDRHPAACASARINV 461
RESULT 6
DCHTCC
ornithine decarboxylase (EC 4.1.1.17) - Chinese hamster
C/Species: Cricetus griseus (Chinese hamster)
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
C/Accession: S09574; A27379
R/Grens, A.; Steglich, C.; Pilz, R.; Scheffler, I.E.
Nucleic Acids Res. 17, 10497, 1989
A/Title: Nucleotide sequence of the Chinese hamster ornithine decarboxylase gene.
A/Reference number: S09574; MUID:90098890; PMID:2602162
A/Accession: S09574
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-455 <GEB>
A/Cross-references: EMBL:X16910; NID:g49439; PIDN:CAA34784.1; PID:g49440
R/Srinivasan, P.R.; Tonin, P.N.; Wensing, E.J.; Lewis, W.H.
J. Biol. Chem. 262, 12871-12878, 1987
A/Title: The gene for ornithine decarboxylase is co-amplified in hydroxyurea-resistant he
A/Reference number: A27379; MUID:87308323; PMID:2867574
A/Accession: A27379
A/Molecule type: mRNA

A;Residues: 163-295, 'EOP', 296-451, 'R', 453-455 <SRI>
 A;Cross-references: GB:U02813
 C;Superfamily: ornithine decarboxylase
 C;Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F;67/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F;354/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 13.3%; Score 200; DB 1; Length 455;
 Best Local Similarity 28.7%; Pred. No. 1,7e-08;
 Matches 76; Conservative 34; Mismatches 71; Indels 84; Gaps 12;

QY 90 PEGYVQ-----ERTIFPKDGNVKTAEVKEGTLVNRILKG-----IDFKEDGN 136
 DB 202 PETFVQALSDARCVF---DMG-----TEVGS-----WTLDIGGFGPGESEPTKLFEBITS 250
 QY 137 ILGHKLE-YNYNHNVYIMADKQNGIKVNFKIRHNI-----EDGSVQLADH 182
 DB 251 VINPALDKYPPPSGVRVIAEPGRYYASAFTLAVNIIAKKIVSKGSDDESSSQTFMY 310
 QY 183 YQONTPIGDGVLLPDNHYLSTQSLSKDPNEK-----RDHNV----- 220
 DB 311 YVNDGVYGSFNCILYDHAHVKPLLPKRPKPEKXSSISWGPTODGLDRIYERCNLPENH 370
 QY 221 -----LLE-----FVTAAGITLG-----MDLYKKL-SHGPPPEVEODDGT 256
 DB 371 VGDWMLFENMGAYTVAAASTFNGFQRPRIYYVMSRPMWQIMKQIQNHGPFPEVEODVGT 430
 QY 257 LPMSCAQSGGMDRHPAACAASRINV 281
 DB 431 LPTSCAQSGGMDRHPAACAASINIV 455

RESULT 7
 DCHDO
 ornithine decarboxylase (EC 4.1.1.17) - human
 N;Alternate names: ODC
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C;Accession: S06900; A33388; J00792; A26957; S14539; A46555; I38077
 R;Van Steeg, H.; van Oostrom, C.T.M.; Martens, J.W.M.; van Kreyl, C.F.; Schepens, J.; W.
 Nucleic Acids Res. 17, 8855-8856, 1989
 A;Title: Nucleotide sequence of the human ornithine decarboxylase gene.
 A;Reference number: S06900; MUID:90067851; PMID:2587220
 A;Accession: S06900
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-461 <VAN>
 A;Cross-references: EMBL:X16277; NID:g35137; PIDN:CAA34353.1; PID:g236667
 R;Fitzgerald, M.C.; Flanagan, M.A.
 DNA 8, 623-634, 1989
 A;Title: Characterization and sequence analysis of the human ornithine decarboxylase gen
 A;Reference number: A33388; MUID:90126232; PMID:2693021
 A;Accession: A33388
 A;Molecule type: DNA
 A;Residues: 1-461 <FIT>
 A;Cross-references: GB:M01740; NID:g189370; PIDN:AA5967.1; PID:g189371; GB:M01061; NID
 R;Hickok, N.J.; Wallofs, J.; Crozat, A.; HalmeKloe, M.; Alhonen, L.; Jaenm, J.; Jaenm
 Gene 93, 257-263, 1990
 A;Title: Human ornithine decarboxylase-encoding loci: nucleotide sequence of the express
 A;Reference number: J00792; MUID:91033036; PMID:2227439
 A;Accession: J00792
 A;Molecule type: DNA
 A;Residues: 1-461 <HIC1>
 A;Cross-references: GB:M3764; NID:g338279; PIDN:AA60564.1; PID:g338280
 R;Hickok, N.J.; Seppanen, P.J.; Gunsalus, G.L.; Jaenm, O.A.
 DNA 6, 179-187, 1987
 A;Title: Complete amino acid sequence of human ornithine decarboxylase deduced from comp
 A;Reference number: A26957; MUID:87246067; PMID:3595418
 A;Accession: A26957
 A;Molecule type: mRNA
 A;Residues: 1-461 <HIC2>
 A;Cross-references: GB:M16650
 R;Flanagan, M.A.; Streng, K.A.; Wagner III, R.L.

submitted to the EMBL Data Library, November 1990
 A;Description: Nucleotide sequence of a human ornithine decarboxylase cDNA.
 A;Reference number: S14539
 A;Accession: S14539
 A;Molecule type: mRNA
 A;Residues: 1-461 <FLA>
 A;Cross-references: EMBL:X55362; NID:g35135; PIDN:CAA39047.1; PID:g35136
 R;Kaczmarek, L.; Calabretta, B.; Ferrati, S.; de Riel, J.K.
 J. Cell. Physiol. 132, 545-551, 1987
 A;Title: Cell-cycle-dependent expression of human ornithine decarboxylase.
 A;Reference number: A46555; MUID:88007889; PMID:3308908
 A;Accession: A46555
 A;Molecule type: mRNA
 A;Residues: 112-461 <KAC>
 A;Cross-references: GB:M20372; NID:g189372; PIDN:AA5968.1; PID:g189373
 R;Helen, U.T.; Deming, M.F.; Heidel, S.M.; Verma, A.K.
 Cancer Res. 50, 2239-2244, 1990
 A;Title: Expression of human chromosome 2 ornithine decarboxylase gene in ornithine decar
 A;Reference number: I38077; MUID:90199754; PMID:2317811
 A;Accession: I38077
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-22 <HS1>
 A;Cross-references: EMBL:X53271; NID:g288102; PIDN:CAA37369.1; PID:g288103
 C;Genetics:
 A;Gene: GDB:ODC1
 A;Cross-references: GDB:119462; OMIM:165640
 A;Map position: 2p25-2p25
 A;Intons: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
 C;Function:
 A;Description: catalyzes the decarboxylation of ornithine to putrescine
 A;Pathway: polyamine biosynthesis
 A;Note: the first and rate-limiting reaction in the pathway
 C;Superfamily: ornithine decarboxylase
 C;Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis; I
 F;69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F;360/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 10.3%; Score 154.5; DB 1; Length 461;
 Best Local Similarity 27.2%; Pred. No. 8.7e-05;
 Matches 72; Conservative 29; Mismatches 83; Indels 81; Gaps 12;

QY 90 PEGYVQ-----ERTIFPKDGNVKTAEVKE-----GDTLVNRILEGIDPKEDGNILG 139
 DB 205 PETFVQALSDARCVF---DMG-----AYGFSMYLIDIGCGPSSDVA-LKFEETGVIN 256
 QY 140 HKLE-YNYNHNVYIMADKQNGIKVNFKIRHNI-----EDGSVQLADH 182
 DB 257 PALDKYPPPSDGVRIIAEPGRYYASAFTLAVNIIAKKIVLKEQTSDDDESSSQTFMY 316
 QY 183 YQONTPIGDGVLLPDNHYLSTQSLSKDPNEK-----RDHNV----- 220
 DB 317 YVNDGVYGSFNCILYDHAHVKPLLPKRPKPEKXSSISWGPTODGLDRIYERCDLPENH 376
 QY 221 -----LLE-----FVTAAGITLG-----MDLYKKL-S-----HGPPPEVEODDGT 256
 DB 377 VGDWMLFENMGAYTVAAASTFNGFQRPRIYYVMSRPMWQIMKQIQNHGPFPEVEODVGT 436
 QY 257 LPMSCAQSGGMDRHPAACAASRINV 281
 DB 437 LPVSCAWEISGMDRHPAACAASINIV 461

RESULT 8
 DCHDO
 ornithine decarboxylase (EC 4.1.1.17) - chicken (fragment)
 C;Species: Gallus gallus (chicken)
 C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C;Accession: A48386; S19892
 R;Johnson, R.; Bulfield, G.
 Anim. Genet. 23, 403-409, 1992
 A;Title: Molecular cloning and sequence analysis of a chicken ornithine decarboxylase cDN
 A;Reference number: A48386; MUID:93036582; PMID:1416246

A:Accession: A48386
 A:Molecule type: mRNA
 A:Residues: 1-450 <002>
 A:Cross-References: EMBL:X64710; NID:963712; PIDN:CAA5965.1; PID:963713
 A:Experimental source: embryos
 A:Note: sequence extracted from NCBI backbone (NCBIP:115105)
 C:Function:
 A:Description: This enzyme catalyzes the decarboxylation of ornithine to putrescine, the C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:5/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:350/Active site: Cys (shared with dimeric partner) #status predicted

Query Match
 Best Local Similarity 7.4%; Score 111.5; DB 1; Length 450;
 Matches 24; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

245 PPEVEEDDGTLPMSQSGMDRHPACASARINV 281
 415 FLAEVEEDVASLPLSCAESGIE-YPAACASASINV 450

RESULT 9

A:Accession: A43563
 A:Molecule type: mRNA
 A:Residues: 1-460 <BAS>
 A:Cross-References: GB:X56316; NID:964953; PIDN:CAA39760.1; PID:964954
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match
 Best Local Similarity 6.9%; Score 103; DB 2; Length 460;
 Matches 64; Conservative 34; Mismatches 81; Indels 90; Gaps 14;

90 PEGVVO---BRTFFKDDGNYKTRAEVFEEDTLVNRILEKG-----IDPKEDGN 136
 205 PGTVOAVSDARCVF---DMG-----AELGFN---MHLLDGGGPPGSEDDVKTKEEITS 253
 137 ILGKLE-YNNNSHNYITMAOKKNGIKVNF-----KIRHNIEDGVQVAD----- 181
 254 VINPALDYPADSGVKLIAEPGRYYVASFTLAVNITAKIVWVEQSGSDDEPAANDK 313
 182 ---HYQMTPTGIDGCVLLEPDNHYLSTGSLSKDPNEK-----RDHWV----- 220
 314 TLMYYVDGVGSGNCLIFDAHAKPVITKPKPEDEKYSSTINGPTCDGIDRIVEREL 373
 221 -----ILE-----FVTAAGITLG---MDELYKKS-----HGFPPEVEEQ 252
 374 PELQGVDMWLFENMGAYVAASFTENGFRPLTYVYVSRPMQIMHDIKKGILPEV--P 431
 253 DDGTLPMSCQSGMDRHPACASARINV 281
 432 DLALHVSQAGSGEMELAPAVCTASINV 460

RESULT 10

A:Accession: H72228
 A:Molecule type: protein
 A:Residues: 1-253
 A:Cross-References: EMBL:X64710; NID:963712; PIDN:CAA5965.1; PID:963713
 A:Experimental source: embryos
 A:Note: sequence extracted from NCBI backbone (NCBIP:115105)
 C:Function:
 A:Description: This enzyme catalyzes the decarboxylation of ornithine to putrescine, the C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:5/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:350/Active site: Cys (shared with dimeric partner) #status predicted

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A12200; MUID:99287316; PMID:10360571
 A:Accession: H72228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-785 <ARN>
 A:Cross-References: GB:AE001806; GB:AE000512; NID:94982196; PIDN:AA036691.1; PID:9498219
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1624

Query Match
 Best Local Similarity 6.7%; Score 101; DB 2; Length 785;
 Matches 40; Conservative 27; Mismatches 64; Indels 60; Gaps 5;

3 SKGERLFTGVVPIVELDGVNGHKFSVSGEGDAGTGLTIKFTCTGKLPVPMPTLV 62
 15 NEGRSFEGTVPGVQAD-----LVKGLIHPHYGSM- 46
 63 TTLTYGVOCFSRRYPDHMKQHDFFKSAMEGVYQERTIFFKDDGNYKTRAEVFEEDTLV 122
 47 -----NEDLFKIEDREWITREFFEFKEDVKEGERVDLVFEGVDTLIS 88
 123 RIELKIDPKEDGNILGKLENNNSHNYITMAOKKNGIKVNFKIRHNIEDGVQVADH 182
 89 DVLANGVTL---GSTEDMFIRPDVINVL---KENHILKVIK-----SPIRVEKT 134
 183 YQNTPTGIDGP 193
 135 LEQNYVGLGCP 145

RESULT 11

ERMUHN
 coatomer complex alpha chain homolog - human
 A:Alternate names: HEP-COP; xenopsin homolog; xenopsin-related peptide precursor
 N:Contains: xenin 25; xenopsin-related peptide

C:Species: Homo sapiens (man)
 C:Date: 10-May-1996 #sequence-revision 08-Nov-1996 #text_change 21-Jul-2000
 C:Accession: J04668; A44317
 R:Chow, V.T.K.; Quek, H.H.
 Gene 169, 223-227, 1996
 A:Title: HEP-COP, a novel human gene whose product is highly homologous to the alpha-sub
 A:Reference number: J04668; MUID:96194806; PMID:8647451
 A:Accession: J04668

A:Molecule type: mRNA
 A:Residues: 1-1224 <CHO>
 A:Cross-References: GB:U24105; NID:91638873; PIDN:AA070879.1; PID:91002369
 A:Experimental source: Hep3b hepatocellular carcinoma cell
 R:Feurle, G.E.; Hamscher, G.; Kusiek, R.; Meyer, H.E.; Metzger, J.W.
 J. Biol. Chem. 267, 22305-22309, 1992
 A:Title: Identification of xenin, a xenopsin-related peptide, in the human gastric mucose
 A:Reference number: A44317; MUID:93054515; PMID:1429581
 A:Accession: A44317
 A:Molecule type: protein
 A:Residues: 1-25 <FEU>
 A:Experimental source: gastric mucosa
 A:Note: sequence extracted from NCBI backbone (NCBIP:117018)
 A:Note: plasma levels of xenin 25, as determined by immunoassay, rise after meals and it
 C:Genetics:
 A:Gene: GDB:COA; HEP-COP

A:Cross-References: GDB:4642787; OMIM:601924
 A:Map position: 1q23-1q25
 C:Superfamily: yeast coatomer complex alpha chain; WD repeat homology
 C:Keywords: duplication; hormone; plasma; stomach
 F:1-25/Product: xenin 25 #status experimental <XNP>
 F:5-38/Domain: WD repeat homology <WD1>
 F:17-25/Product: xenopsin-related peptide #status predicted <XRP>
 F:47-80/Domain: WD repeat homology <WD2>
 F:89-122/Domain: WD repeat homology <WD3>

F:131-164/Domain: WD repeat homology <WD4>
F:201-234/Domain: WD repeat homology <WD5>
F:245-278/Domain: WD repeat homology <WD6>

Query Match

Best Local Similarity 26.1%; Pred. No. 9.2; Length 1224;
Matches 47; Conservative 28; Mismatches 68; Indels 37; Gaps 8;

QY 91 EGVYQERTTFPKDGNKYTRAEVKEDTLVNRLE-LKGDIPKEDGILGKLEVNNSH 149

Db 632 KGYEVALHFKDE---KTRFSLALECCNIEIALMAKALDDKCKEGLVALLQGNHQ 688

QY 150 NYVIMADKOKNGIKVKN-----XIRNIEDGSVQ--LADHYQONTPIGGPVLLPD 198

Db 689 IVEKCYQRTNPKDPKVSFLYLITGNLEKTRKMKIAEIRKMSGHYQVALYLD----- 741

QY 199 NHYLSIQSLSKDPNEKRDHVLLEFVTAAGITLGMELKXLSHGPPEVVEKDDGTLF 258

Db 742 ---VSEKRVRLIKKCGCKS-----LAYLTLA--THGLDEKASLKEFTDPKKE-----TTP 786

RESULT 12

T20721
hypothetical protein F25F2.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-May-2000

C:Accession: T20721; T21343; T23842

R:Ainscough, R.

submitted to the EMBL Data Library, August 1994

A:Accession: T20721

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4307 <M12>

A:Cross-references: EMBL:Z35662; PIDN:CAA84721.1; GSPDB:GN00021; CESP:F25F2.2

A:Experimental source: clone F10G11

R:Ainscough, R.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19410

A:Accession: T21343

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4307 <M12>

A:Cross-references: EMBL:Z35599; PIDN:CAA84661.1; GSPDB:GN00021; CESP:F25F2.2

A:Experimental source: clone F25F2

R:Subston, T.

submitted to the EMBL Data Library, June 1994

A:Reference number: Z19806

A:Accession: T23842

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

QY 101 FKD-----DGNVKT-RAEVKESDPLVNRLELKGDIPKEDGILGKLEVN 146

Db 3433 DEBEHQNGILRGVLESGISQEPRAQVQSRSTHLPVDPNTGDIWSD-----HSTQGL 3487

QY 147 NSHVYIMADKOKNGIK-----VNEKIR----- 169

Db 3488 HFEVNTV-IDSKEVTVSVYEVHTSIDNDVIDHAVISIRIRMSVDPEFMKHYKEPRRIS 3546

QY 170 HNI---EDGSVQLADHYQONTPIGGPVLLPDNHYSIQSLSKDPNEKRDHVLLEFV 226

Db 3547 HHNLNDDSSIQLI--VQAVP-----STES-----ERRSRNSMEDVEILM 3586

QY 227 AAGITLGM-----DELYKLSHGF 245

Db 3587 TIAQGLGREGVLYKPDHITYSRLKNDP 3610

RESULT 13

KXRTCI

proprotein convertase 1 (EC 3.4.21.93) precursor - rat

N:Alternate names: furin homolog PC1; kexin homolog PC1; prohormone cleavage enzyme; prot

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-May-2000

C:Accession: A41556; S27361; S36358

R:Bloomquist, B.T.; Ripper, B.A.; Maine, R.E.

Mol. Endocrinol. 5, 2014-2024, 1991

A:Reference number: A41556; MUID:92168040; PMID:1791845

A:Accession: A41556

A:Molecule type: mRNA

A:Residues: 1-752 <BLD>

A:Cross-references: GB:M76705; NID:G203508; PIDN:AAA40945.1; PID:G203509

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

Endocrinology 129, 3053-3063, 1991

A:Title: Isolation of two complementary deoxyribonucleic acid clones from a rat insulinoma

to endocrine and neuroendocrine tissues in rats.

A:Reference number: S27361; MUID:92063860; PMID:1954888

A:Accession: S27361

A:Molecule type: mRNA

A:Residues: 1-513, 'A', 515-752 <HAK>

A:Cross-references: EMBL:M83745

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

submitted to the EMBL Data Library, February 1992

A:Reference number: S36358

A:Accession: S36358

A:Molecule type: mRNA

A:Residues: 1-366, 'T', 367-513, 'A', 515-752 <HAK>

QY 18 ELDGVDVNGHK-----FSVSGEGEDATGKILKRICITGKLP-----VWPVTLVITLTYG 68

Db 540 ELDGVDVNGHK-----FSVSGEGEDATGKILKRICITGKLP-----VWPVTLVITLTYG 68

QY 69 VQCSRPYDHNKQDFFKSNAPPEGYVQERTTFPKDGNKYTRAEVKEDTLVNRLE 128

Db 594 T---SSQPEHMKQ-----PRVYT-----SYNTVQNDRRGVEKGVVAVVEKP 631

QY 129 IDFFEDGNILGKLEVNNSHVYIMADKOKNGIKVKNFKIRNIEDGSVQLADHYQONT 188

Db 129 IDFFEDGNILGKLEVNNSHVYIMADKOKNGIKVKNFKIRNIEDGSVQLADHYQONT 188

QY 227 AAGITLGM-----DELYKLSHGF 245

Db 3587 TIAQGLGREGVLYKPDHITYSRLKNDP 3610

RESULT 13

KXRTCI

proprotein convertase 1 (EC 3.4.21.93) precursor - rat

N:Alternate names: furin homolog PC1; kexin homolog PC1; prohormone cleavage enzyme; prot

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-May-2000

C:Accession: A41556; S27361; S36358

R:Bloomquist, B.T.; Ripper, B.A.; Maine, R.E.

Mol. Endocrinol. 5, 2014-2024, 1991

A:Reference number: A41556; MUID:92168040; PMID:1791845

A:Accession: A41556

A:Molecule type: mRNA

A:Residues: 1-752 <BLD>

A:Cross-references: GB:M76705; NID:G203508; PIDN:AAA40945.1; PID:G203509

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

Endocrinology 129, 3053-3063, 1991

A:Title: Isolation of two complementary deoxyribonucleic acid clones from a rat insulinoma

to endocrine and neuroendocrine tissues in rats.

A:Reference number: S27361; MUID:92063860; PMID:1954888

A:Accession: S27361

A:Molecule type: mRNA

A:Residues: 1-366, 'T', 367-513, 'A', 515-752 <HAK>

A:Cross-references: EMBL:M83745; NID:G205062; PIDN:AAA41476.1; PID:G205063

C:Comment: This protein is a member of a family of subtilisin-like proteinases responsible

C:Comment: This protein lacks a classical hydrophobic transmembrane segment but may assoc

C:Keywords: glycoprotein; hydrolase; serine proteinase

F:125-110/Domain: signal sequence #status predicted <SIG>

F:111-752/Product: prohormone-processing proteinase PC1 #status predicted <MAT>

F:158-396/Domain: subtilisin homology <SBT>

QY 18 ELDGVDVNGHK-----FSVSGEGEDATGKILKRICITGKLP-----VWPVTLVITLTYG 68

Db 540 ELDGVDVNGHK-----FSVSGEGEDATGKILKRICITGKLP-----VWPVTLVITLTYG 68

QY 69 VQCSRPYDHNKQDFFKSNAPPEGYVQERTTFPKDGNKYTRAEVKEDTLVNRLE 128

Db 594 T---SSQPEHMKQ-----PRVYT-----SYNTVQNDRRGVEKGVVAVVEKP 631

QY 129 IDFFEDGNILGKLEVNNSHVYIMADKOKNGIKVKNFKIRNIEDGSVQLADHYQONT 188

Db 129 IDFFEDGNILGKLEVNNSHVYIMADKOKNGIKVKNFKIRNIEDGSVQLADHYQONT 188

QY 227 AAGITLGM-----DELYKLSHGF 245

Db 3587 TIAQGLGREGVLYKPDHITYSRLKNDP 3610

RESULT 13

KXRTCI

proprotein convertase 1 (EC 3.4.21.93) precursor - rat

N:Alternate names: furin homolog PC1; kexin homolog PC1; prohormone cleavage enzyme; prot

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-May-2000

C:Accession: A41556; S27361; S36358

R:Bloomquist, B.T.; Ripper, B.A.; Maine, R.E.

Mol. Endocrinol. 5, 2014-2024, 1991

A:Reference number: A41556; MUID:92168040; PMID:1791845

A:Accession: A41556

A:Molecule type: mRNA

A:Residues: 1-752 <BLD>

A:Cross-references: GB:M76705; NID:G203508; PIDN:AAA40945.1; PID:G203509

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

Endocrinology 129, 3053-3063, 1991

A:Title: Isolation of two complementary deoxyribonucleic acid clones from a rat insulinoma

to endocrine and neuroendocrine tissues in rats.

A:Reference number: S27361; MUID:92063860; PMID:1954888

A:Accession: S27361

A:Molecule type: mRNA

A:Residues: 1-366, 'T', 367-513, 'A', 515-752 <HAK>

A:Cross-references: EMBL:M83745; NID:G205062; PIDN:AAA41476.1; PID:G205063

C:Comment: This protein is a member of a family of subtilisin-like proteinases responsible

C:Comment: This protein lacks a classical hydrophobic transmembrane segment but may assoc

C:Keywords: glycoprotein; hydrolase; serine proteinase

F:125-110/Domain: signal sequence #status predicted <SIG>

F:111-752/Product: prohormone-processing proteinase PC1 #status predicted <MAT>

F:158-396/Domain: subtilisin homology <SBT>

Db 632 TONSINGNLVVK--NSSSSSVEDRRDECVQCAPSKAMR-----LLQSAFSEKNT 680
 QY 189 IGDGPVLLPDNHYLSTQSAISKDPNEKRDPHMLLEFVTAAGITLGMDELTKLSHGFPPE 248
 Db 681 -----SKQS--SKTSPAKLS-----VPYEGLYEALFKLNK-----PSQ 711
 QY 249 VERQDD 254
 Db 712 LEDSED 717

RESULT 14

C64468
 hypotetrical protein M01348 - Methanococcus jannaschii
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #ext_change 21-Jul-2000
 C/Accession: C64468
 R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kitzness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Kleink, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A/Reference number: A64300; M01D:96337993; PMID:8688087
 A/Accession: C64468
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-336 <BU>
 A/Cross-references: GB:U67574; GB:L77117; NID:gl591978; PIDN:AB99360.1; PID:gl591990; T
 C/Genetics:
 A/Map position: FOR1295121-1296131
 A/Start codon: TTG

Query Match 6.1%; Score 92; DB 2; Length 336;
 Best Local Similarity 24.9%; Pred. No. 6.6;
 Matches 66; Conservative 37; Mismatches 80; Indels 82; Gaps 17;

QY 20 DGDVNGHFSVSGE---GEGDAYGKL--TKFTCTTGKLPVPW-----PTL 61
 Db 91 DGDY---YNLSGELSTASIFAKIGKIDITGNFNSGSE---WYNDVTKDANSEDL 142
 QY 62 VTTLTYGV-----QCFERYDPMKQ-----HDFKSAPEGYQERTIFPKDDG 105
 Db 143 KSVLTFRDSYERKEILNREP-HLKKLFEDNIYNFNSDFPFMM-----MEFIGAG 192
 QY 106 NYKTRAEVKEF-----GDTLVNRIELKIDFKEDGNIL--GHKLEYNVNSHNYI 153
 Db 193 NMRKFLVEYERFKKIKSCQISNEIYNEI-IRPD-KMSDLIAHLIKENYEKCLYV 250
 QY 154 MADK-----QKNGIVNFKIRHNIDGSVQLADHYQONTPIGDGPVLLP-DNH 201
 Db 251 MLFKEYFDLDEFNEIKKKIMLYDIAYNLKNGVKKKEWMLRNLNEYKEIKRPLPTY 310
 QY 202 LSTQSALSKD-PNEKRDPHMLLEFY 225
 Db 311 ---KDAHNDLLNELIDLYVYVKEFI 332

RESULT 15

G81355
 tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jejuni
 C/Species: Campylobacter jejuni
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C/Accession: G81355
 R/Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling,
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVlier, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A/Reference number: A81250; M01D:20150912; PMID:10688204
 A/Accession: G81355
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-357 <PAR>

A/Cross-references: GB:AL139076; GB:AL111168; NID:96968128; PIDN:CA873096.1; PID:9696827
 A/Experimental source: serotype O2, strain NCTC 11168
 C/Genetics:
 A/Name: tRNA, Cj0831c
 C/Keywords: methyltransferase; S-adenosylmethionine

Query Match 6.1%; Score 91; DB 2; Length 357;
 Best Local Similarity 19.5%; Pred. No. 8.6;
 Matches 53; Conservative 43; Mismatches 84; Indels 92; Gaps 11;

QY 80 KOHDFEKSAMPEGYVQERTIFPKDGNKYTRAVER--EGDTLV----- 121
 Db 14 EKHGFTKXKPEFFYKDPKLFASKDKHRTAELSFYHNDTLFYAMDPSKKKXITLEY 73
 QY 122 -----NRIELK--GIDFKEDGNILGHKLELYNN----- 147
 Db 74 LDFADEKICAFMPRLLEYLRQDNKLEKLFGEVFFLTQKQELSTLLYHKNLEIDIKSNLEN 133
 QY 148 -SHNYIMADQKNGIKYNEKIRH-----NIEDGSVQLADHYQONTPIGDGPVLLPDNHY 201
 Db 134 LSNILHNLIRSKGKGLIFETENLRQTLNTQDRKI---FYEFN-----NDCF 178
 QY 202 LSTQSALSKDPNEK-----RDHMLLEFVTAAG-ITLGMDELTKLSHGFPPEV 249
 Db 179 IQPNTAL---NEKMITWVCSILNTQKMDLLELYCGYGNFTLALAPFFKI---LATBI 231
 QY 250 EEODDGTLPMSCAQSGMDRPPACASARINV 281
 Db 232 SKSNINFALKNCELNNTTNIHFARLSSEELSL 263

Search completed: May 14, 2004, 08:27:17
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 ; Search time 17 Seconds

(without alignments)
860.689 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500

Sequence: 1 MYSKGEELFTGVPIIVELD.....AQESGMDRHPAACAARINLV 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1251	83.4	238	1	PA2212 aequorea vi
2	221.5	14.6	461	1	P00860 mus musculi
3	210.5	14.0	461	1	P27113 mus pahari
4	208.5	13.9	461	1	P09057 rattus norv
5	200	13.3	455	1	P14019 cricetus norv
6	154.5	10.3	461	1	P11926 homo sapien
7	142.5	9.5	461	1	P27117 bos taurus
8	111.5	7.4	450	1	P27118 gallus galli
9	103	6.9	460	1	P27120 xenopus lae
10	99.5	6.6	1224	1	P36121 homo sapien
11	96.5	6.4	1224	1	P36121 homo sapien
12	92.5	6.2	752	1	P27954 bos taurus
13	92	6.1	336	1	P28840 rattus norv
14	91	6.1	357	1	P28840 rattus norv
15	91	6.1	357	1	P28840 rattus norv
16	90.5	6.0	861	1	P27113 mus pahari
17	90.5	6.0	861	1	P27113 mus pahari
18	89	5.9	439	1	P37870 bacillus su
19	88.5	5.9	879	1	P24506 discodyce o
20	88.5	5.9	866	1	P24506 discodyce o
21	88.5	5.9	866	1	P24506 discodyce o
22	87.5	5.8	6632	1	UN89_CAREL
23	87.5	5.8	6632	1	UN89_CAREL
24	86.5	5.8	689	1	AC2L_HUMAN
25	86.5	5.8	689	1	AC2L_HUMAN
26	86.5	5.8	504	1	YC03_XLEPN
27	86.5	5.8	658	1	ADAS_CAVPO
28	86.5	5.8	860	1	SYL_SALT
29	86.5	5.8	860	1	SYL_SALT
30	86.5	5.8	860	1	SYL_SALT
31	86.5	5.8	1603	1	VIT4_CAREL
32	85.5	5.7	788	1	DPOE_YEAST
33	85.5	5.7	879	1	SYL_XYLEFT

34	85.5	5.7	1164	1	BAG_STRAG
35	85	5.7	466	1	Y065_MYCGE
36	85	5.7	3744	1	TRAI_YEAST
37	84.5	5.6	484	1	RGS9_BOVIN
38	84.5	5.6	613	1	PEPF_MYCPU
39	84.5	5.6	795	1	D152_HAIRIN
40	84.5	5.6	797	1	D151_HAIRIN
41	83.5	5.6	793	1	D153_HAIRIN
42	83.5	5.6	859	1	SET7_HUMAN
43	83	5.5	366	1	SET7_HUMAN
44	83	5.5	1214	1	BRF1_HUMAN
45	83	5.5	1585	1	P3K3_DICDI

ALIGNMENTS

RESULT 1
GFP_AEQVI STANDARD; PRT; 238 AA.

AC PA2212; Q17104; Q27903.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Green fluorescent protein.

GN GFP.

OS Aequorea victoria (Jellyfish).

CC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Leptomedusae;

CC Aequoreidae; Aequorea.

OX NCBI_TaxID=6100;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RA Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,

RA Cormier M.J.;

RT "Primary structure of the Aequorea victoria green-fluorescent

protein.";

RL Gene 111:1229-233 (1992).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=94185810; PubMed=8137953;

RA Inouye S., Tsuji F.I.;

RT "Aequorea green fluorescent protein. Expression of the gene and

fluorescence characteristics of the recombinant protein.";

RL FEBS Lett. 341:277-280 (1994).

RN [3]

RP SEQUENCE FROM N.A.

RA MEDLINE=9729832; PubMed=9154981;

RA Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;

RT "Enhanced expression in tobacco of the gene encoding green fluorescent

protein by modification of its codon usage.";

RL Plant Mol. Biol. 33:989-999 (1997).

RN [4]

RP CHROMOPHORE.

RA MEDLINE=93192221; PubMed=8448132;

RA Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;

RT "Chemical structure of the hexapeptide chromophore of the Aequorea

green-fluorescent protein.";

RL Biochemistry 32:1212-1218 (1993).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RA MEDLINE=9635565; PubMed=8703075;

RA Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,

RT "Crystal structure of the Aequorea victoria green fluorescent

protein.";

RL Science 273:1392-1395 (1996).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RA MEDLINE=98294543; PubMed=9631087;

RA Yang F., Moss L.G., Phillips G.N. Jr.;

RT "The molecular structure of green fluorescent protein.";

RL Nat. Biotechnol. 14:1246-1251 (1996).

[7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
 RX MEDLINE=98455509; PubMed=9782051;
 RA Wachter R.M., Elsliiger M.A., Kallio K., Hanson G.T., Remington S.J.;
 RT "Structural basis of spectral shifts in the yellow-emission variants
 RT of green fluorescent protein.";
 RL Structure 6:1267-1277(1998).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=99338303; PubMed=10220315;
 RA Elsliiger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
 RT "Structural and spectral response of green fluorescent protein
 RT variants to changes in pH.";
 RL Biochemistry 38:5296-5301(1999).
 CC -1- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
 CC blue chemiluminescence of the protein aequorin into green
 CC fluorescence by energy transfer. Fluoresces in vivo upon
 CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
 CC Absorbs light maximally at 395 nm and exhibits a smaller
 CC absorbance peak at 470 nm. The fluorescence emission spectrum
 CC peaks at 509 nm with a shoulder at 540 nm.
 CC -1- SUBUNIT: Monomer.
 CC -1- TISSUE SPECIFICITY: Photocytes.
 CC -1- PTM: Contains a covalently attached chromophore, which is composed
 CC of modified amino acid residues. The chromophore is formed upon
 CC cyclization of the residues Ser-dehydrotyr-Gly.
 CC -1- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
 CC chimeric proteins of GFP linked to other proteins where it
 CC functions as a fluorescent protein tag. GFP tolerates N- and C-
 CC terminal fusion to a broad variety of proteins. It has been
 CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
 CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
 CC marker in living cells, it allows for a wide range of applications
 CC where it may function as a cell lineage tracer, reporter of gene
 CC expression, or as a measure of protein-protein interactions.
 CC -1- DATABASE: NAME-Protein Spotlight;
 CC NOTE-Issue 11 of June 2001;
 CC WWW="http://www.expasy.org/spotlight/articles/spclt011.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M62654; AAA27722.1; -;
 DR EMBL; M62653; AAA27721.1; -;
 DR EMBL; L29345; AAA58246.1; -;
 DR EMBL; X96418; CAA65278.1; -;
 DR PIR; J50692; J01514.
 DR PDB; 1B9C; 17-NOV-00.
 DR PDB; 1BFP; 07-JUL-97.
 DR PDB; 1CAF; 14-JUN-00.
 DR PDB; 1EMA; 08-NOV-96.
 DR PDB; 1EMB; 16-JUN-97.
 DR PDB; 1EMC; 20-AUG-97.
 DR PDB; 1EME; 20-AUG-97.
 DR PDB; 1EMF; 20-AUG-97.
 DR PDB; 1EMK; 12-MAY-99.
 DR PDB; 1EML; 20-AUG-97.
 DR PDB; 1EMM; 20-AUG-97.
 DR PDB; 1F03; 17-NOV-00.
 DR PDB; 1F0B; 17-NOV-00.
 DR PDB; 1GFL; 11-JAN-97.
 DR PDB; 1HCU; 15-JAN-97.
 DR PDB; 1HUY; 04-JUL-02.
 DR PDB; 1JBY; 07-JUN-03.
 DR PDB; 1JBY; 07-JUN-03.
 DR PDB; 1KBS; 28-AUG-02.
 DR PDB; 1KYP; 10-APR-02.

DR PDB; 1KXR; 10-APR-02.
 DR PDB; 1KXS; 10-APR-02.
 DR PDB; 1YFP; 28-OCT-98.
 DR PDB; 2EMD; 20-AUG-97.
 DR PDB; 2EMN; 20-AUG-97.
 DR PDB; 2EMO; 20-AUG-97.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPUNSCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 KW Luminescence; 3D-structure.
 FT CROSSLINK 65 67 5-imidazolinone (Ser-Gly).
 FT MOD RES 66 66 F->Y.
 FT VARIANT 100 100 F->S.
 FT VARIANT 108 108 T->S.
 FT VARIANT 141 141 L->M.
 FT VARIANT 219 219 V->I.
 FT CONFLICT 2 2 S->G (IN REF. 3).
 FT CONFLICT 25 25 H->Q (IN REF. 2).
 FT CONFLICT 80 80 Q->R (IN REF. 3).
 FT CONFLICT 157 157 Q->P (IN REF. 2).
 FT CONFLICT 172 172 E->K (IN REF. 2).
 FT HELIX 4 8
 FT STRAND 12 22
 FT TURN 23 24
 FT STRAND 25 36
 FT TURN 37 40
 FT STRAND 41 48
 FT TURN 49 50
 FT HELIX 57 60
 FT TURN 61 63
 FT HELIX 69 71
 FT STRAND 73 73
 FT HELIX 76 81
 FT HELIX 83 86
 FT TURN 87 90
 FT STRAND 92 100
 FT TURN 101 102
 FT STRAND 105 115
 FT TURN 116 117
 FT STRAND 118 128
 FT TURN 132 133
 FT STRAND 135 139
 FT TURN 141 141
 FT STRAND 148 155
 FT TURN 156 159
 FT STRAND 160 171
 FT TURN 172 173
 FT STRAND 176 187
 FT STRAND 199 208
 FT TURN 211 212
 FT STRAND 217 227
 SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;
 Query Match 83.4%; Score 1251; DB 1; Length 238;
 Best Local Similarity 98.3%; Pred. No. 6.2e-94;
 Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VSKGEELFTGVVPIVHLDQVNGHKRSVSGEBSGDATYKLTLPKICTTGKLPVMPPTL 61
 DB 1 MSKGEELFTGVVPIVHLDQVNGHKRSVSGEBSGDATYKLTLPKICTTGKLPVMPPTL 60
 QY 62 VTLTYGVQCFSRYPHMKQHPFKSAMPGYQERTIFPKDGNKTRAEVFEQDTLV 121
 DB 61 VTTFSYGVQCFSRYPHMKQHPFKSAMPGYQERTIFPKDGNKTRAEVFEQDTLV 120
 QY 122 NRTELKGDPEKDNILGHKLRYNNSHNYIMADKQKGIKVNFKIRANIEDGSVQLAD 181
 DB 121 NRTELKGDPEKDNILGHKLRYNNSHNYIMADKQKGIKVNFKIRANIEDGSVQLAD 180
 QY 182 HYQONPPIGDGPTLPPNHYLSTQSALSKDPNKRBDHMTLEFVTAAGTILGMDELTK 239


```

FT TURN 60 61
FT STRAND 62 67
FT HELIX 68 70
FT HELIX 74 83
FT TURN 84 84
FT STRAND 86 89
FT HELIX 92 100
FT TURN 101 102
FT HELIX 105 107
FT STRAND 108 110
FT HELIX 117 125
FT TURN 126 127
FT STRAND 130 133
FT HELIX 136 145
FT TURN 147 148
FT STRAND 150 155
FT STRAND 172 172
FT HELIX 174 186
FT TURN 187 188
FT STRAND 190 195
FT TURN 205 205
FT HELIX 206 225
FT TURN 226 226
FT STRAND 231 233
FT HELIX 248 252
FT STRAND 255 257
FT TURN 268 273
FT STRAND 270 273
FT HELIX 277 280
FT HELIX 281 283
FT STRAND 284 296
FT STRAND 313 318
FT TURN 322 326
FT HELIX 327 331
FT STRAND 339 340
FT TURN 346 347
FT STRAND 350 356
FT TURN 362 363
FT STRAND 365 373
FT TURN 377 378
FT STRAND 380 383
FT HELIX 391 393
FT HELIX 397 399
FT STRAND 404 410
FT HELIX 411 417
SQ SEQUENCE 461 AA; 5163 MW; DEF1753FA2ACA0A30 CRC64;

Query Match 14.8%; Score 221.5; DB 1; Length 461;
Best Local Similarity 87.8%; Pred. No. 1.7e-10;
Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 234 MDLYKKT-SHGPPPEVEEDDGLTFMSCAQSSEGMDRHPAACAARINV 281
DB 413 MWOLMKQIQSHGPPPEVEEDDGLTFMSCAQSSEGMDRHPAACAARINV 461

RESULT 3
DCOR MUSPA STANDARD; PRT; 461 AA.
AC P27119;
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Mus pahari (Shrew mouse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250571; PubMed=1374399;
RA Johannes G., Berger F.G.;

```

```

RT "Alterations in mRNA translation as a mechanism for the modification
RT of enzyme synthesis during evolution. The ornithine decarboxylase
RT model.";
RL J. Biol. Chem. 267:10108-10115 (1992).
CC -1- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1- SUBUNIT: Homodimer.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M87223; AAA9847.1; -.
DR PIR; I55356; I55356.
DR HSSP; P00860; 7ODC.
DR MED; MGI:97402; Odc.
DR InterPro; IPR000183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
DR PRINTS; PR01179; ODADCBXLABS.
DR PROSITE; PS00878; ORN_DC_2_1; 1.
DR PROSITE; PS00879; ORN_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW phosphorylation.
FT BINDING 69 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT SITE 360 360 BY SIMILARITY.
FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
FT (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51203 MW; 61ABCBEC7BFB92 CRC64;

Query Match 14.0%; Score 210.5; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 1.3e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;

QY 90 PEGYVQ---ERTTFKDDGNYKTRAEYKFRGDTLVNRIELKG-----IDPKDGN 136
DB 205 PETFQAVSDARCVF--DMG-----TEVGF--MYLLDIGGPPGSEDYTKLFEELITS 253
QY 137 ILGHKLE-YYNNSHNYVTMADKKMGKIKVNFKIRHNI-----EDGSYQL 179
DB 254 VINPALDKTFPSDSGVRLIAEPGRYVASAFTLVAVITAKTWTKEQPSDEDSNBO 313
QY 180 ADHYQANTPIGDGPVTLDPNHYLSTQSALSKDPNEK-----RDHMY----- 220
DB 314 FMYYVNDGVYSPNCLIVDHAHVAKALLKRPKPEKYYSSSIWPTCGDLRIYRCMLP 373
QY 221 -----LLE---FTTAAGITLG-----MDLYKKT-SHGPPPEVEED 253
DB 374 ENHWGDWMLFENMGAYTYAAASTFNGFORPNIIYYMSRPMWQIMKRIOHSGPPEVEED 433
QY 254 DGLTFMSCAQSSEGMDRHPAACAARINV 281
DB 434 DGLTFMSCAQSSEGMDRHPAACAARINV 461

RESULT 4
DCOR RAT STANDARD; PRT; 461 AA.
AC P09057;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Liver;
RX MEDLINE=88335556; PubMed=3419906;
RA van Steeg H., van Oostrom C.T.M., van Kranen H.J., van Kreijl C.F.;
RT "Nucleotide sequence of the rat ornithine decarboxylase gene.";
RL Nucleic Acids Res. 16:8173-8174 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86167817; PubMed=3443298;
RA van Kranen H.J., van de Zande L., van Kreijl C.F., Bisschop A.,
RT "Cloning and nucleotide sequence of rat ornithine decarboxylase
RT cDNA.";
RL Gene 60:145-155 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=89255378; PubMed=2722815;
RA Wen L., Huang J.K., Blackshear P.J.;
RT "Rat ornithine decarboxylase gene. Nucleotide sequence, potential
RT regulatory elements, and comparison to the mouse gene.";
RL J. Biol. Chem. 264:9016-9021 (1989).
CC -1 CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1 COFACTOR: Pyridoxal phosphate.
CC -1 PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1 SUBUNIT: Homodimer.
CC -1 SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X07944; CAA30765.1; -
DR EMBL, M16982; AAA41737.1; -
DR EMBL, J04792; AAA66286.1; ALT_SEQ.
DR EMBL, J04791; AAA66164.1; -
DR PIR, A27361; DCRTD.
DR HSP, P00860; 7ODC.
DR InterPro; IPR00183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR PRINTS; PR01179; ODADCBXLASE.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW Phosphorylation.
FT BINDING 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 360 BY SIMILARITY.
FT MOD_RES 303 PHOSPHORYLATION (BY CK2)
FT (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51047 MW; CEF9D268DDDD0D6 CRC64;
Query Match 13.9%; Score 208.5; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 1.9e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGVQ---ERTFFKDDGNYKRAEYKFGDTLVNRIELKG-----IDFKEDGN 136
DB 205 PETTVQVSPDARCVF--DMG-----TEVGFSS---MYLLDGGGPGSGEDTKLFEETIS 253
QY 137 ILGHKE-YVNSHNVTIMADKONGIKVNFIRHNI-----EDGSVOL 179
DB 254 VINALKKTPPSDGVATIAFGKRYVASAFTLAVNIATKWTKEQTGSDEDSNQOT 313

```

```

QY 180 ADHYQNTPIGDGVILLPDNHYLSTGSLSKDPNEK-----RDHW----- 220
DB 314 LMTYVNDGVYSGENCILYDAHVAFALLQKRPKDEKYSSTWGPCTDGLDRIVERCSLP 373
QY 221 -----LLE-----FTVAAGITL-----GMDELYKKL-SHGFPPEVEED 253
DB 374 EMHVGDMMLTENMCAIVVAALSTFNGFRPNITYYVMSRSMQMLKQIQSHGFPPEVEED 433
QY 254 DGTTPMSCAQSQSGMDRHPAACASARINV 281
DB 434 VGTLPMSCAQSQSGMDRHPAACASASINV 461
RESULT 5
DCOR CRIGR STANDARD; PRT; 455 AA.
AC P14019;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9098890; PubMed=2602162;
RA Grens A., Steglich C., Pilz R., Scheffler I.E.;
RT "Nucleotide sequence of the Chinese hamster ornithine decarboxylase
RL Nucleic Acids Res. 17:10497-10497 (1989).
RN [2]
RP SEQUENCE OF 163-455 FROM N.A.
RX MEDLINE=87308329; PubMed=2887574;
RA Strinvaasen P.R., Tonin P.N., Wensing E.J., Lewis W.H.;
RT "The gene for ornithine decarboxylase is co-amplified in hydroxyurea-
RT resistant hamster cells.";
RL J. Biol. Chem. 262:12871-12878 (1987).
CC -1 CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1 COFACTOR: Pyridoxal phosphate.
CC -1 PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1 SUBUNIT: Homodimer.
CC -1 SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X16910; CAA34784.1; -
DR EMBL, J02813; AAA36963.1; -
DR PIR, S09574; DCHYOC.
DR HSP, P00860; 7ODC.
DR InterPro; IPR00183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR PRINTS; PR01179; ODADCBXLASE.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW Phosphorylation.
FT BINDING 67 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 354 BY SIMILARITY.
FT MOD_RES 297 PHOSPHORYLATION (BY CK2)

```

```

FT      CONFLICT      452      452      (BY SIMILARITY)
FT      SEQUENCE      455 AA; 50453 MW; 14162985F18C89F CRC64;
SQ      SEQUENCE      455 AA; 50453 MW; 14162985F18C89F CRC64;

Query Match      13.3%; Score 200; DB 1; Length 455;
Best Local Similarity 28.7%; Pred. No. 9.2e-09;
Matches 76; Conservative 34; Mismatches 71; Indels 84; Gaps 12;

QY      PEGYVQ-----ERTFFKDGNYKTRAEVFEEDTTLVNRILEKG-----IDFEKDGAN 136
DB      202 PEFYQALSDACV--DMG-----TEVGF--MYLLIDGGFPGSDTLKKEEITS 250
QY      137 ILGHKLE-YNVNSHNVYTMADKXKNGIKYKFKIRNI-----EDGSYQLADH 182
DB      251 VINPALDKFFPPDSGVRVIAEFGRYVASAFTLAVNIILAKIIVSKSDDEDESSQTFMY 310
QY      183 YQGNTPIGGVPVLPDNNHLSQSAISKDQNEK-----RDHMV----- 220
DB      311 YVNDGYGSENCILTDHAKVFKLPPRRKPKDEKYYSSSTWGTCDGLDIVERCNLPENH 370
QY      221 -----LLE-----FYTAAGITLG-----MDLYKKL-SHGFPPEVERODDGT 256
DB      371 VGDWMLFEMMGAYTYTAAASTFNGFQRPSTIYVMSRPMMQIQNHGFPPEVEEDVGT 430
QY      257 LPMSCAOGSGMRRHPAACASARINV 281
DB      431 LPISCAOGSGMDRHPAACASASINV 455

RESULT 6
DCOR HUMAN
ID      DCOR HUMAN      STANDARD;      PRT;      461 AA.
AC      P11926;
DT      01-OCT-1989 (Rel. 12, Created)
DT      01-OCT-1989 (Rel. 12, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN      ODC1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID:9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90067851; PubMed=2587220;
RA      van Steeg H., van Oostrom C.T.M., Martens J.W.M., van Kreyl C.F.,
RA      Schepens J., Wieringa B.;
RT      "Nucleotide sequence of the human ornithine decarboxylase gene.";
RL      Nucleic Acids Res. 17:8855-8856(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9012632; PubMed=2693021;
RA      Fitzgerald M.C., Flanagan M.A.;
RT      "Characterization and sequence analysis of the human ornithine
RT      decarboxylase gene.";
RL      DNA 8:623-634(1989).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87246067; PubMed=3595418;
RA      Hickok N.J., Seppanen P.J., Gunsalus G.L., Jaenne O.A.;
RT      "Complete amino acid sequence of human ornithine decarboxylase
RT      deduced from complementary DNA.";
RL      DNA 6:179-187(1987).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91033036; PubMed=2227439;
RA      Hickok N.J., Wahlfors J., Crozat A., Halmekytoe M., Alhonen L.,
RA      Jaenne O.A.;
RT      "Human ornithine decarboxylase-encoding loci: nucleotide sequence of
RT      the expressed gene and characterization of a pseudogene.";
RL      Gene 93:257-263(1990).
RN      [5]
RP      SEQUENCE FROM N.A.

```

```

RX      MEDLINE=92285144; PubMed=1598217;
RA      Moshier J.A., Osborne D.L., Skunca M., Dosescu J., Gilbert J.D.,
RA      Fitzgerald M.C., Polidori G., Wagner R.L., Fritzen Degen S.J.,
RA      Luk G.D., Flanagan M.A.;
RT      "Multiple promoter elements govern expression of the human ornithine
RT      decarboxylase gene in colon carcinoma cells.";
RL      Nucleic Acids Res. 20:2581-2590(1992).
RN      [6]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90202959; PubMed=2318872;
RA      Moshier J.A., Gilbert J.D., Skunca M., Dosescu J., Almodovar K.M.,
RA      Luk G.D.;
RT      "Isolation and expression of a human ornithine decarboxylase gene.";
RL      J. Biol. Chem. 265:4884-4892(1990).
RN      [7]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22386257; PubMed=12477932;
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Ditchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ueda T.B., Toshimuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska M., Smallus D.E.,
RA      Sutterch A., Schein J.R., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [8]
RP      SEQUENCE OF 1-22 FROM N.A.
RX      MEDLINE=90199754; PubMed=2317811;
RA      Hsieh J.T., Denning M.F., Heide S.M., Verna A.K.;
RT      "Expression of human chromosome 2 ornithine decarboxylase gene in
RT      ornithine decarboxylase-deficient Chinese hamster ovary cells.";
RL      Cancer Res. 50:2239-2244(1990).
RN      [9]
RP      X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX      MEDLINE=20090911; PubMed=10623504;
RA      Almud J.U., Oliveira M.A., Kern A.D., Grishin N.V.,
RA      Phillips M.A., Hacker M.L.;
RT      "Crystal structure of human ornithine decarboxylase at 2.1 A
RT      resolution: structural insights to antilyme binding.";
RL      J. Mol. Biol. 295:7-16(2000).
CC      - CARBOXYLIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC      - CORFACTOR: Pyridoxal phosphate.
CC      - PATHWAY: Polyamine biosynthesis, first (rate-limiting) step.
CC      - SUBUNIT: Homodimer.
CC      - SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC      DECARBOXYLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; M31061; AAA60563.1; -
DR      EMBL; M16650; AAA59966.1; -
DR      EMBL; X16277; CAA34353.1; -
DR      EMBL; M33764; AAA60564.1; -
DR      EMBL; X55362; CAA39047.1; -

```

DR EMBL; M81740; AAA59967.1; -.
 DR EMBL; M34158; AAA59969.1; -.
 DR EMBL; BC025296; AAH25296.1; -.
 DR EMBL; X53271; CA37369.1; -.
 DR PIR; S06900; DCHDO.
 DR PDB; 1D7K; 25-OCT-00.
 DR Genew; HGNC:8109; ODC1.
 DR GK; P11926; -.
 DR MM; 165640; -.
 DR GO; GO:0004586; F:ornithine decarboxylase activity; TAS.
 DR GO; GO:0006596; P:polymine biosynthesis; NAS.
 DR InterPro; IPR00183; Decarboxylase.
 DR InterPro; IPR009006; Racem decarbox_C.
 DR Pfam; PF02784; Orn_Arg_dec_N; 1.
 DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
 DR PRINTS; PR01179; ODADECBLASE.
 DR PROSITE; PS00878; ODR_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 DR Lyase; Decarboxylase; Pyridoxal phosphate; Polymine biosynthesis;
 KM Phosphorylation; 3D-structure.
 FT BINDING 69 69 PYRIDOXAL PHOSPHATE.
 FT ACT_SITE 360 360 BY SIMILARITY.
 FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
 (BY SIMILARITY).
 Q -> E (IN REF. 3).
 FT CONFLICT 415 415
 FT STRAND 11 13
 FT TURN 16 17
 FT HELIX 20 28
 FT TURN 29 30
 FT STRAND 40 44
 FT HELIX 45 58
 FT TURN 60 61
 FT STRAND 62 67
 FT HELIX 68 70
 FT TURN 74 83
 FT TURN 84 84
 FT STRAND 86 89
 FT HELIX 92 99
 FT TURN 100 102
 FT HELIX 105 107
 FT STRAND 108 110
 FT HELIX 117 125
 FT TURN 126 127
 FT STRAND 130 133
 FT HELIX 136 145
 FT TURN 147 148
 FT STRAND 150 155
 FT STRAND 172 172
 FT HELIX 174 186
 FT TURN 187 188
 FT STRAND 190 195
 FT TURN 200 201
 FT TURN 205 205
 FT HELIX 206 225
 FT TURN 226 226
 FT STRAND 231 233
 FT HELIX 248 262
 FT HELIX 265 267
 FT STRAND 270 273
 FT HELIX 277 280
 FT HELIX 281 283
 FT STRAND 284 296
 FT STRAND 313 318
 FT TURN 322 326
 FT HELIX 327 331
 FT STRAND 339 340
 FT TURN 346 347
 FT STRAND 350 356
 FT TURN 362 363
 FT STRAND 365 373
 FT TURN 377 378
 FT STRAND 380 383
 FT HELIX 391 393

FT HELIX 397 399
 FT STRAND 404 409
 FT HELIX 411 420
 SQ SEQUENCE 461 AA; 51148 MM; 8CCB8CEB0823C5 CRC64;
 Query Match 10.3%; Score 154.5; DB 1; Length 461;
 Best Local Similarity 27.2%; Pred. No. 4.4e-05;
 Matches 72; Conservative 29; Mismatches 83; Indels 81; Gaps 12;
 QY 90 PEGYVQ---ERITFFEDDOANYTRAAYKE-----GDTLVNRLEKIGDFEDNGNITLG 139
 DB 205 PEFVQAISDARCVF--DWG-----AEVGFMYLIDIGGFFPSSEBYK-LKFEIIGVIN 256
 QY 140 HKLE-YVNSHNYVIMADKXGKIKVNFKIRHNI-----EDSSVQADH 182
 DB 257 PALDKYPPSSGVRITIAEPKRYVAAFTLVNIIAKKIVLKQTSDDDESSGQTFMY 316
 QY 183 YQONTPIGDGPVLLPNNHYLSTQSALSKEPNEK-----RDENV----- 220
 DB 317 YVNDGVYGFNCILYDHAHVKPLQKRPKDEKYSSSIWPTCDGLDRIVERCDLPEMH 376
 QY 221 ----LIE---FVTAAGITLG---MDELYKLS-----HGPPPYEBODDGT 256
 DB 377 VGDMLFENNGAYTAAASTFNGFORPTIYVWSGPANQMOQFONPDPPEVEODPAST 436
 QY 257 LPMSCAQESGMDRHPAACAASARINV 281
 DB 437 LPVSCAWESGMRHRAACASASINV 461
 RESULT 7
 DCOB BOVIN
 ID DCOR BOVIN STANDARD; PRT; 461 AA.
 AC P2717;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
 OS ODC1 OR ODC.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 ON NCBI_Taxid=9913;
 RX MEDLINE=95293216; Pubmed=7774801;
 RA Yao J., Zadworny D., Kuhnlein U., Hayes J.F.;
 RT "Molecular cloning of a bovine ornithine decarboxylase cDNA and its
 use in the detection of restriction fragment length polymorphisms in
 Holsteins";
 RL Genome 38:325-331(1995).
 CC -|- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
 CC -|- COFACTOR: Pyridoxal phosphate.
 CC -|- PATHWAY: Polymine biosynthesis; first (rate-limiting) step.
 CC -|- SUBUNIT: Homodimer.
 CC -|- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
 DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M92441; AAA92339.1; -.
 DR EMBL; U36394; AAA79849.1; -.
 DR EMBL; U18531; AAA6696.1; -.
 DR HSSP; P11926; 1D7K
 DR InterPro; IPR00183; Decarboxylase2.

DR InterPro; IPR009006; Racem decarbox_C.
 DR Pfam; PF02784; Orn Arg dec N; 1.
 DR Pfam; PF02784; Orn DAP Arg dec; 1.
 DR PRINTS; PR01179; ODAOCRXYLASE.
 DR PROSITE; PS00878; ORN_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
 KW Phosphorylation.
 FT BINDING 69 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT ACT_SITE 360 360 BY SIMILARITY.
 FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
 FT (BY SIMILARITY).
 SQ SEQUENCE 461 AA; 51345 MW; 486098433868FA CRC64;
 Query Match 9.5%; Score 142.5; DB 1; Length 461;
 Best Local Similarity 27.2%; Pred. No. 0.00041;
 Matches 72; Conservative 32; Mismatches 80; Indels 81; Gaps 15;
 QY 90 PEGYVQ---ERTFPKDDGNKTRAEVKE-----GDTLVNRIELKGDPEKEDNIIIG 139
 DB 205 PEFVQALSDARCFE---DMG-----AEVGFNNYLLDIGGFGPSBDVK-LKFEITSVIN 256
 QY 140 HKLE YVNSNHYVMADKQKGIKVKRINIEDGSVQ-----ADHYQNTF----- 188
 DB 257 PALDKYFPGSDGVRILAEGRYVASAFLLAVNIITAKLVKEQTSDEESTDRTMY 316
 QY 189 -IGDG-----PVL-----LPDNHYSTQ-----SALSKDPNEKRDH 218
 DB 317 YNDGVVYSGFNCILVDHNVKPLQKRPKDSKYSSSIWPTCGDLRIYRCMLPEMH 376
 QY 219 M---VLLE---FTVTAAGITLG---MDELYKLS-----HGPPPEVEEQDGT 256
 DB 377 VGDWMLFEMMGAYTVAASTFGFQRPRTIYYWMSGPTQMQOIRTODFPPEVEEPDVG 436
 QY 257 LPMSCAOSGMDRHPACASARIN 281
 DB 437 LPVSCAMSGMRHSAACSTRIN 461
 RESULT 8
 DCOOR CHICK STANDARD; PRT; 450 AA.
 AC P27118;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ornithine decarboxylase (EC 4.1.1.17) (ODC) (Fragment).
 OS ODC.
 GN Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RK MEDLINE=93036582; PubMed=1416246;
 RA Johnson R., Bullfield G.;
 RT "Molecular cloning and sequence analysis of a chicken ornithine
 RT decarboxylase cDNA."
 RT Anim. Genet. 23:403-409(1992).
 CC -1- CATABOLIC ACTIVITY: L-ornithine = putrescine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
 CC DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X64710; CAA45965.1; -.
 DR PIR; A48386; DCCHO.
 DR HSSP; P11926; 10TK.
 DR InterPro; IPR000183; Decarboxylase2.
 DR InterPro; IPR009006; Racem decarbox_C.
 DR Pfam; PF02784; Orn Arg dec N; 1.
 DR Pfam; PF02784; Orn DAP Arg dec; 1.
 DR PRINTS; PR01179; ODAOCRXYLASE.
 DR PROSITE; PS00878; ORN_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
 KW Phosphorylation.
 FT NON_TER 1 1
 FT BINDING 59 59 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT ACT_SITE 350 350 BY SIMILARITY.
 FT MOD_RES 293 293 PHOSPHORYLATION (BY CK2)
 FT (BY SIMILARITY).
 SQ SEQUENCE 450 AA; 49734 MW; 5EAC75DF3D17AD2C CRC64;
 Query Match 7.4%; Score 111.5; DB 1; Length 450;
 Best Local Similarity 64.9%; Pred. No. 0.13; Indels 1; Gaps 1;
 Matches 24; Conservative 5; Mismatches 7;
 QY 245 FPPEVEEQDGTLPMSCAOSGMDRHPACASARIN 281
 DB 415 FLAVEEQDVASTPLSCACSGIE-YRATCSASIN 450
 RESULT 9
 DCOOR XENLA STANDARD; PRT; 460 AA.
 AC P27120;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ornithine decarboxylase 1 (EC 4.1.1.17) (XODC1).
 GN ODC1 OR ODC.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISUE=Oocyte;
 RK MEDLINE=91209251; PubMed=2088731;
 RA Bassez T., Paris J., Omtali F., Dorel C., Osborne H.B.;
 RT "Post-transcriptional regulation of ornithine decarboxylase in
 RT Xenopus laevis oocytes."
 RT Development 110:955-962(1990).
 RN [2]
 RP DEVELOPMENTAL EXPRESSION.
 RC MEDLINE=92104181; PubMed=1761057;
 RK Osborne H.B., Duval C., Ghoda L., Omtali F., Bassez T., Coffino P.;
 RA "Expression and post-transcriptional regulation of ornithine
 RT decarboxylase during early Xenopus development."
 RT Eur. J. Biochem. 202:575-581(1991).
 CC -1- CATABOLIC ACTIVITY: L-ornithine = putrescine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- ENZYME REGULATION: REGULATED POST-TRANSLATIONALLY BY POLYAMINES.
 CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
 CC -1- DEVELOPMENTAL STAGE: LEVELS INCREASE IN THE EMBRYOS 1.5-2 HOURS
 CC AFTER FERTILIZATION AND REACH A MAXIMUM AT 6 HOURS POST-
 CC FERTILIZATION, THEN DECREASE TO A VERY LOW LEVEL IN EARLY
 CC GASTRULATING EMBRYOS (12 HRS POST-FERTILIZATION).
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
 CC DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X56316; CA93760.1; -
 CC PIR; A43563; A43563.
 CC HSSP; P00860; 70DC.
 CC InterPro; IPR000183; Decarboxylase2.
 CC InterPro; IPR000906; Racem decarbox_C.
 CC Pfam; PF02784; Orn_Arg_dec_N; 1.
 CC Pfam; PF00278; Orn_DAP_Arg_dec; 1.
 CC PRINTS; PR01179; ODADCRXLASE.
 CC PROSITE; PS00878; ODR_DC_2_1; 1.
 CC PROSITE; PS00879; ODR_DC_2_2; 1.
 CC Lysase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
 KW phosphorylation.
 FT BINDING 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT ACT_SITE 361 361 BY SIMILARITY.
 FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
 FT (BY SIMILARITY).
 SQ SEQUENCE 460 AA; 50830 MW; CCGF87FD165877EE CRC64;
 Query Match 6.9%; Score 103; DB 1; Length 460;
 Best Local Similarity 23.8%; Pred. No. 0.63;
 Matches 64; Conservative 34; Mismatches 81; Indels 90; Gaps 14;
 QY 90 PEGYVQ---ERTIFKDDGNYKRAVEKFGDITVRIELKG-----IDFKENGN 136
 Db 205 PQTVAQVSDARCVF--DMG-----AELGFN---MHLLDGGGPGSEVDYKLFEEITIS 253
 QY 137 ILHGKLE-YNNYNSHNVYIMADKOKNGIKVNF-----KIRHNIEDGSYQLAD----- 181
 Db 254 VINPALDKYFPADSGVKILAEPRGYVASSFTLAVNIILAKVMNKGSGSDDEEDANDK 313
 QY 182 ---HYQNTPTIGDGPVILPNHNVYSTGSAKSDPENK-----RDHNV----- 220
 Db 314 TLMVYVDGYYGVSFNCILFDHAHYVPVLTKKPKDEKYSISSIWPTCDGLDRIYERPEL 373
 QY 221 -----LLE-----FVTAAGITLG---MDELYKKS-----HGFPPEVVEQ 252
 Db 374 PELQGVDMLEFENMGAYVVAALASTFNGQRPPTLYVNSRPHQMMDIKHGIILPEV--P 431
 QY 253 DGGTLPMSCAQSGMDRHPAACASARINV 281
 Db 432 DLSALHVSCAQSGMELAPVCTAASINV 460
 RESULT 10
 COPA_HUMAN STANDARD; PRT; 1224 AA.
 ID COPA_HUMAN
 AC P53621;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coatomer alpha subunit (Alpha-coat protein) (Alpha-COP) (HEPCOP)
 DE (HEP-COP) [Contains: Xenin (Xenopsin-related peptide); Proxenin].
 GN COPA.
 OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCI_TaxID=9606;
 RX MEDLINE=96194806; PubMed=8647451;
 RA Chow V.T.K., Quek H.H.;
 RT "HEP-COP, a novel human gene whose product is highly homologous to
 the alpha-subunit of the yeast coatomer protein complex.";
 RL Gene 169:223-227(1996).
 RN [2]
 RP SEQUENCE OF 1-25 (XENIN).
 RC TISSUE=Gastric mucosa;

RX MEDLINE=93054515; PubMed=1429581;
 RA Feurle G.E., Hamscher G., Kusiek R., Meyer H.E., Metzger J.W.;
 RT "Identification of xenin, a xenopsin-related peptide, in the human
 gastric mucosa and its effect on exocrine pancreatic secretion.";
 RL J. Biol. Chem. 267:22305-22309(1992).
 RN [3]
 RP PROCESSING OF COPA TO PRODUCE XENIN.
 RX MEDLINE=98032493; PubMed=9365789;
 RA Chow V.T., Quek H.H.;
 RT "Alpha coat protein COPA (HEP-COP): presence of an Alu repeat in cDNA
 and identity of the amino terminus to xenin.";
 RL Ann. Hum. Genet. 61:369-373(1997).
 RN [4]
 RP REVIEW ON XENIN.
 RX MEDLINE=98192336; PubMed=9533652;
 RA Feurle G.E.;
 RT "Xenin -- a review.";
 RL Peptides 19:609-615(1998).
 CC -1- FUNCTION: The coatomer is a cytosolic protein complex that binds
 to lysine motifs and reversibly associates with Golgi non-
 clathrin-coated vesicles, which further mediate biosynthetic
 protein transport from the ER, via the Golgi up to the trans Golgi
 network. Coatomer complex is required for budding from Golgi
 membranes, and is essential for the retrograde Golgi-to-ER
 transport of dilysine-tagged proteins. In mammals, the coatomer
 can only be recruited by membranes associated to ADP-ribosylation
 factors (ARFs), which are small GTP-binding proteins; the complex
 also influences the Golgi structural integrity, as well as the
 processing, activity, and endocytic recycling of LDL receptors (By
 similarity).
 CC -1- FUNCTION: XENIN STIMULATES EXOCRINE PANCREATIC SECRETION. IT
 INHIBITS PENTAGASTRIN-STIMULATED SECRETION OF ACID, TO INDUCE
 EXOCRINE PANCREATIC SECRETION AND TO AFFECT SMALL AND LARGE
 INTESTINAL MOTILITY. IN THE GUT, XENIN INTERACTS WITH THE
 NEUTROPHILIN RECEPTOR.
 CC -1- SUBUNIT: Oligomeric complex that consists of at least the alpha,
 beta, beta', gamma, delta, epsilon and zeta subunits.
 CC -1- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
 ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
 VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY). XENIN IS
 SECRETED.
 CC -1- TISSUE SPECIFICITY: UNIFORMLY EXPRESSED IN A WIDE RANGE OF ADULT
 AND FETAL TISSUES. XENIN IS FOUND IN GASTRIC, DUODENAL AND JEJUNAL
 MUCOSA. CIRCULATES IN THE BLOOD. SEEMS TO BE CONFINED TO SPECIFIC
 ENDOCRINE CELLS.
 CC -1- DEVELOPMENTAL STAGE: XENIN IS RELEASED INTO THE CIRCULATION AFTER
 A MEAL.
 CC -1- SIMILARITY: Contains 6 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; U24105; AAB70879.1; -
 CC PIR; J04668; ERHDAH.
 CC Genew; HENC:2230; COPA.
 CC MIM; 601924; -
 CC GO; GO:0005803; C:secretory vesicle; TAS.
 CC GO; GO:0005215; F:transporter activity; TAS.
 CC GO; GO:0006888; P:ER to Golgi transport; TAS.
 CC InterPro; IPR006692; Coatomer WDAD.
 CC InterPro; IPR008941; TPR-like.
 CC InterPro; IPR001880; WD40.
 CC Pfam; PF04053; Coatomer WDAD; 1.
 CC Pfam; PF00400; WD40; 6.
 CC PRINTS; PR00320; GPROTEINRPT.
 CC PRODOM; PD000018; WD40; 5.
 CC SMART; SMO0320; WD40; 7.
 CC PROSITE; PS00678; WD_REPEATS_1; 1.

DR PROSITE; PS50082; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Transports; Protein transport; Golgi stack; Endoplasmic reticulum;
 KM Membrane; Repeat; WD repeat; Hormone.
 FT PEPTIDE 1 25 XENIN.
 FT REPEAT 1 35 PRORENIN.
 FT REPEAT 7 37 WD 1.
 FT REPEAT 49 79 WD 2.
 FT REPEAT 91 121 WD 3.
 FT REPEAT 133 163 WD 4.
 FT REPEAT 203 233 WD 5.
 FT REPEAT 247 277 WD 6.
 SQ SEQUENCE 1224 AA; 138331 MW; 5A8E8684D58F155D CRC64;
 Query Match 6.6%; Score 99.5; DB 1; Length 1224;
 Best Local Similarity 26.1%; Pred. No. 4;
 Matches 47; Conservative 28; Mismatches 68; Indels 37; Gaps 8;
 QY 91 EGYVQERTFFPDGNGYKTRAEVKEGDTLVNRIR-LKGDIFKEDGNILGHKLEYNNSH 149
 DB 632 KGYEVALHFKYDE---KTRFSIALBEGCNIEIALBAKALDDKNCWEMKGEVALIQGNHQ 688
 QY 150 NYIYIMADKQKNGIKYVNF-----KIRHNIEDGSVQ--LADHYQONTPIGSGVLLPD 198
 DB 689 IYEMCYQRTKNDKXSYLYLITGNIEKLRKMMKIAIRKMSGHYQNALYIGD----- 741
 QY 199 NHYLSQSALSKDNEKRDHNVLEFYTAGITLGMDELYKKLSHGPPVEBQDDGTLF 258
 DB 742 ---VSEKRYILKNGGQS-----LAVLTAA--THGLDEBAESLKTPTDPEK-----TTP 786
 RESULT 11
 COPA_BOVIN
 ID Q27954; STANDARD; PRT; 1224 AA.
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coatomer alpha subunit (Alpha-coat protein) (Alpha-COP) (HEPCOP)
 DE (HEP-COP) [Contains: Xenin (Xenopsin-related peptide); Proxenin].
 GN Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN SEQUENCE FROM N.A.
 RA Wieland F.I.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBS databases.
 CC -!- FUNCTION: The coatomer is a cytosolic protein complex that binds
 to dilysine motifs and reversibly associates with Golgi non-
 clathrin-coated vesicles, which further mediate biosynthetic
 protein transport from the ER, via the Golgi up to the trans Golgi
 network. Coatomer complex is required for budding from Golgi
 membranes, and is essential for the retrograde Golgi-to-ER
 transport of dilysine-tagged proteins. In mammals, the coatomer
 can only be recruited by membranes associated to ADP-ribosylation
 factors (ARFs), which are small GTP-binding proteins; the complex
 also influences the Golgi structural integrity, as well as the
 processing, activity, and endocytic recycling of LDL receptors (By
 similarity).
 CC -!- FUNCTION: XENIN STIMULATES EXOCRINE PANCREATIC SECRETION. IT
 INHIBITS PENTAGASTRIN-STIMULATED SECRETION OF ACID, TO INDUCE
 EXOCRINE PANCREATIC SECRETION AND TO AFFECT SMALL AND LARGE
 INTESTINAL MOTILITY. IN THE GUT, XENIN INTERACTS WITH THE
 NEUROTENSIN RECEPTOR (BY SIMILARITY).
 CC -!- SUBUNIT: Oligomeric complex that consists of at least the alpha,
 beta, beta', gamma, delta, epsilon and zeta subunits.
 CC SUBCELLULAR LOCATION: THE COATOMER IS CYTOSOLIC OR POLYMERIZED
 ON THE CYTOSOLIC SIDE OF THE GOLGI, AS WELL AS ON THE
 VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY). XENIN IS
 SECRETED (BY SIMILARITY).

CC -!- SIMILARITY: Contains 6 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 DR EMBL; X96768; CAA65543.1; -
 DR InterPro; IPR006692; Coatomer_WDAD.
 DR InterPro; IPR008941; TPR-1like.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF04053; Coatomer_WDAD; 1.
 DR Pfam; PF00400; WD40; 6.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR PRODOM; PD000018; WD40; 5.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; 1.
 DR PROSITE; PS50082; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Transports; Protein transport; Golgi stack; Endoplasmic reticulum;
 KM Membrane; Repeat; WD repeat; Hormone.
 FT PEPTIDE 1 25 XENIN.
 FT REPEAT 1 35 PRORENIN.
 FT REPEAT 7 37 WD 1.
 FT REPEAT 49 79 WD 2.
 FT REPEAT 91 121 WD 3.
 FT REPEAT 133 163 WD 4.
 FT REPEAT 203 233 WD 5.
 FT REPEAT 247 277 WD 6.
 SQ SEQUENCE 1224 AA; 138358 MW; C9350BF2AC00683D CRC64;
 Query Match 6.4%; Score 96.5; DB 1; Length 1224;
 Best Local Similarity 25.6%; Pred. No. 7;
 Matches 46; Conservative 29; Mismatches 68; Indels 37; Gaps 8;
 QY 91 EGYVQERTFFPDGNGYKTRAEVKEGDTLVNRIR-LKGDIFKEDGNILGHKLEYNNSH 149
 DB 632 KGYEVALHFKYDE---KTRFSIALBEGCNIEIALBAKALDDKNCWEMKGEVALIQGNHQ 688
 QY 150 NYIYIMADKQKNGIKYVNF-----KIRHNIEDGSVQ--LADHYQONTPIGSGVLLPD 198
 DB 689 IYEMCYQRTKNDKXSYLYLITGNIEKLRKMMKIAIRKMSGHYQNALYIGD----- 741
 QY 199 NHYLSQSALSKDNEKRDHNVLEFYTAGITLGMDELYKKLSHGPPVEBQDDGTLF 258
 DB 742 ---VSEKRYILKNGGQS-----LAVLTAA--THGLDEBAESLKTPTDPEK-----TTP 786
 RESULT 12
 NECL_RAT
 ID P28840; STANDARD; PRT; 752 AA.
 AC 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neuroendocrine convertase 1 precursor (EC 3.4.21.93) (NEC 1) (PCL)
 DE (Prohormone convertase 1) (Proprotein convertase 1).
 GN PCCK1 OR NECL OR NEC-1 OR BDP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92168040; Pubmed=1791845;
 RA Bloomquist B.T., Eipper B.A., Mains R.E.;
 RT "Prohormone-converting enzymes: regulation and evaluation of function
 using antisense RNA."
 RL Mol. Endocrinol. 5:2014-2024(1991).
 RN [2]


```

RP SEQUENCE FROM N.A.
RX MEDLINE=92063860; PubMed=1954888;
RA Hakes D.J., Birch N.P., Mezey A., Dixon J.E.;
RT "Isolation of two complementary deoxyribonucleic acid clones from a
RT rat insulinoma cell line based on similarities to Kex2 and furin
RT sequences and the specific localization of each transcript to
RT endocrine and neuroendocrine tissues in rats.";
RL Endocrinology 129:3053-3063(1991).
CC -1- FUNCTION: Involved in the processing of hormone and other protein
CC precursors at sites comprised of pairs of basic amino acid
CC residues. Substrates include POMC, renin, enkephalin, dynorphin,
CC somatostatin and insulin.
CC -1- CATALYTIC ACTIVITY: Release of protein hormones, neuropeptides and
CC renin from their precursors, generally by cleavage of -Lys-Arg-|-
CC bonds.
CC -1- COFACTOR: Calcium.
CC -1- SUBCELLULAR LOCATION: Localized in the secretion granules.
CC -1- SIMILARITY: Belongs to peptidase family S8. Furin subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M75705; AAA40945.1; -
DR EMBL: M83745; AAA41476.1; -
DR PIR: A41556; KXRTCL.
DR HSSP: Q45670; IDBI.
DR MEROPS: S08.072; -
DR InterPro: IPR002029; Peptidase_S8.
DR InterPro: IPR002884; Peptidase_S8B.
DR InterPro: IPR009020; Protease_inhib.
DR Pfam: PF01463; P_propioctein; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000717; P_domain; 1.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase; Serine protease; Glycoprotein; Zymogen; Calcium; Signal.
FT SIGNAL 1 27
FT PROPEP 28 110
FT CHAIN 111 752
FT DOMAIN 122 410
FT DOMAIN 739 751
FT ACT_SITE 167 167
FT ACT_SITE 208 208
FT ACT_SITE 382 382
FT CARBOHYD 173 173
FT CARBOHYD 401 401
FT CARBOHYD 645 645
FT CONFLICT 366 366
FT CONFLICT 514 514
SQ SEQUENCE 752 AA; 84120 MW; F630AD830A076DED CRC64;

Query Match 6.2%; Score 92.5; DB 1; Length 752;
Best Local Similarity 24.4%; Pred. No. 8.1;
Matches 60; Conservative 23; Mismatches 80; Indels 77; Gaps 13;
18 ELGDGVNGHK-----FSVSGEGEDATYGLTKLFICTTGKLP-----VPMPTLVTTLYTG 68
Db 540 ERLTSPRGKFNQMDMSVHTMGENV--GTWTLKATLDMSGMONGRIVNKKLL-----HG 593
QY 69 VQCESRPDMKQKDFKSKAMPEGVQERITTFKDDQGNVYTRAEVKEGDTLVNRIELKG 128
Db 594 T-----SSQPEHMKQ-----PRYLT-----SYNTVQDRRGVKKMNVVVEKP 631
QY 129 IDFKEDNIIIGHKLEVYNSHNHYIMADKQNGIKVFKIRHNIEDSVQLADHYOQNTP 188
Db 632 TQNSLNGNLLVPR-----NSSSSVEDRDRDQVQAPSKAMLR-----LLQSAFSAKNP 680

```

```

QY 189 IGDGPVLDPNHYLSTQSAISKDPENKRDHMYLFEVTAAGITAGMDLYKLKLSHGFPPE 248
Db 681 -----SKQS-----SKIPSAKLS-----VPIEGLYALEKLNK-----PSQ 711
QY 249 VERQDD 254
Db 712 LERSED 717

RESULT 13
YD48 METVA STANDARD; PRT; 336 AA.
AC 058743;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ1348.
GN MJ1348.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2461 / ATCC 43067;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
RA Overbeek R., Kirkness E.F., Weisskopf K.G., Merriam J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Ulfteback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
RN [2]
RP DISCUSSION ON SEQUENCE.
RX MEDLINE=98362752; PubMed=9697413;
RA Kyriades N.C., Woese C.R.;
RT "Tetratric-peptide-repeat proteins in the archaeon Methanococcus
RT jannaschii.";
RL Trends Biochem. Sci. 23:245-247(1998).
CC -1- CAUTION: According to Ref.2, this sequences contains TPR repeats.
CC These are not detected using our methodology.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67574; AAB99360.1; -
DR PIR: C64468; C64468.
DR TIGR: M01348; -
DR InterPro: IPR008941; TPR-like.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 336 AA; 40443 MW; 41BF2F28CDB036FE CRC64;

Query Match 6.1%; Score 92; DB 1; Length 336;
Best Local Similarity 24.9%; Pred. No. 3.3;
Matches 66; Conservative 37; Mismatches 80; Indels 82; Gaps 17;
QY 20 DGDVNGHKFSVSGE-----GEGDATYGLT-----TLKFICTTGKLPVPMW-----PTL 61
Db 91 DGDY-----VYLSEGLSTIASIFAKIGKLDITKFNLSGGE-----WYNVYIKDANSEDL 142
QY 62 VTLLTYTG-----QCFSRYPDMKQ-----HDFPKSAMPEGVQERITTFKDDG 105

```

Db 143 KSVLTFEDESYEERKEILNREP-HLRKLFEDNTYNENSDFFEDFM-----MEFISAG 192
 QY 106 NKXLTAEVKE-----GDTLVNRLTELKIDPEKEDNTL--GHKLEYVNSHNHYI 153
 Db 193 NKRRLEVEYEEFKKIKSCQISNEIYNEL-IRFPD-KMSMDLALAHLLKENYECLYIV 250
 QY 154 VADK-----QKNGIKVNFKIRHNIEDGSVQLADHYOONTPIGDGPAVLDP-DNHY 201
 Db 251 MLFKEFYDFDEPNEIEKKNNILIVLAVLKNQNVKKEEMLNRLNEIYKEIKRELPNTY 310
 QY 202 ISTQSALSKPD-PNEKRDHNVLLFEV 225
 Db 311 ---KDAHNDLLEILDYVLAKEFI 332

RESULT 14

TRMA CAMJE STANDARD; PRT; 357 AA.
 ID TRMA CAMJE

AC Q9PP92;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE tRNA (Traci)-5-methyltransferase (BC 2.1.1.35) (tRNA (M-5-US4) - methyltransferase) (RUMT).
 GN TRMA OR C00831C.

OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_Taxid=197;
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parthill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.K., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences."
 RL Nature 403:665-668 (2000).

RT -1- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
 54 (M-5-US4) in all tRNA (by similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing thymine.
 CC -1- SIMILARITY: Belongs to the RNA M5U methyltransferase family. Trna
 CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AL139076; CAB75096.1; -.

DR PIR: G81355; G81355.

DR HAMAP: MF_01011; -; 1.

DR InterPro: IPR000051; SAM bind.

DR InterPro: IPR001566; Trna.

DR PROSITE: PS01230; TRMA_1; 1.

DR PROSITE: PS01231; TRMA_2; FALSE_NEG.

KW Transferase; Methyltransferase; tRNA processing; Complete proteome.

FT DOMAIN 207 213

FT SIMILARITY: S-ADENOSYLMETHIONINE BINDING (BY
 FT SIMILARITY).

FT ACT SITE 315 315

FT SEQUENCE 357 AA; 42276 MW; CEC5328347CEB97 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 357;
 Best Local Similarity 19.5%; P. No. 4.3;
 Matches 53; Conservative 43; Mismatches 84; Indels 92; Gaps 11;

QY 80 KOHDFEFSAMEBEGYVQRTIFFKDDGNYKTRAEVKE--EGDTLV----- 121
 Db 14 EKHSFIKKYFEFFETTKQFKASDKKHYRRAELSFHENDTLFYAMFDPSKKKYLIEY 73
 QY 122 -----NRIEK--GIDPEKEDNTLGHKLEYNN----- 147
 Db 74 LDFADEKICAMPRLLELYLRQDNKLEKLFGEVLETTQKEISTLVLRKNIEDIKSNLEN 133
 QY 148 -SHNVYIMADQKNGKIGVNEKIRH-----NTEDSVQLADHYOONTPIGDGPAVLDPNHY 201
 Db 134 LSNLTHTLTLARSGKGLIFETENVRQTLNTQDKKI-----FYERN-----NDCE 178
 QY 202 ISTQSALSKPDNEK-----RDHNVLLFEVTAAG-ITLGMDLYKKLSHGFPPEV 249
 Db 179 IQPNTAL-----NEWMITWVCEILNTQKRMDELLEYCGVGNFTLALAPFFKI---LATEI 231
 QY 250 EEDDGTTPMSCAQBSGMDRHPAACAARINV 281
 Db 232 SKSNINPALKNCEILNNTNTHPARLSSEELSL 263

RESULT 15

CP51 CANCA STANDARD; PRT; 533 AA.
 ID CP51 CANCA

AC P50859; Q02312;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 51 (Ec 1.14.13.70) (CYP51) (P450-LI1) (Sterol 14-
 DE alpha-demethylase) (lanosterol 14-alpha demethylase) (P450-14DM).
 GN ERG1 OR CYP51.
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_Taxid=5478;
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=2001-L5;
 RX MEDLINE=96161286; PubMed=8593007;
 RA Geber A., Hitchcock C.A., Swartz J.E., Pullen F.S., Maresden K.E.,
 RA Kwon-Chung K.J., Bennett J.E.;
 RA "Deletion of the Candida glabrata ERG3 and ERG1 genes: effect on cell
 RT viability, cell growth, sterol composition, and antifungal
 RT susceptibility."
 RL Antimicrob. Agents Chemother. 39:2708-2717 (1995).
 RN [2]

SEQUENCE OF 60-473 FROM N.A.

RC STRAIN=ATCC 2001;
 RX MEDLINE=95081364; PubMed=7989540;
 RA Burgener-Kalnitz P., Zuber U.P., Jaunin P., Buchman T.G., Bille J.,
 RA Roseler M.;
 RA "Rapid detection and identification of Candida albicans and
 RT Torulopsis (Candida) glabrata in clinical specimens and
 RT species-specific nested PCR amplification of a cytochrome P-450
 RT lanosterol-alpha-demethylase (L141) gene fragment."
 RL J. Clin. Microbiol. 32:1902-1907 (1994).

CC -1- FUNCTION: Catalyzes C14-demethylation of lanosterol which is
 CC critical for ergosterol biosynthesis. It transforms lanosterol
 CC into 4,4'-dimethyl cholesterol-8,14,24-triene-3-beta-ol (By
 CC similarity).

CC -1- CATALYTIC ACTIVITY: Obusufolol + 3 O(2) + 3 NADPH = 4-alpha-
 CC methyl-5-alpha-ergosta-8,14,24(28)-trien-3-beta-ol + formate + 3
 CC NADP(+) + 3 H(2)O.

CC -1- PATHWAY: Ergosterol biosynthesis.

CC -1- SIMILARITY: Belongs to the cytochrome P450 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; L40389; AAB02329.1; -.
DR EMBL; S75389; AAB32679.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme;
KW Sterol biosynthesis; NADP.
FT METAL 472 472 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 64 64 I -> M (IN REF. 2).
FT CONFLICT 473 473 I -> T (IN REF. 2).
SQ SEQUENCE 533 AA; 61305 MW; A0506C17507B6E7 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 533;
Best Local Similarity 20.3%; Pred. No. 7;
Matches 54; Conservative 44; Mismatches 106; Indels 62; Gaps 11,

QY 25 GHKFSV---GEGGDATYGGKLTIKFTCTTGKLPVPMFTLVTLTYGVQCFSRYPDH--M 79
Db 109 GHEFIENAKLADVSAEAAYSHL-----TTPVFGKVIYDCPNHRLM 149
QY 80 KOHDFEKSAM--PEGVY-----OERTIEFKDDGNVYKTRAEVKEGDTLVNRIELKGIDF 131
Db 150 EQKFFVKALYKEAFVRYVPLIAEIIYKFRNSKNFKINENSGIVDWVSQPEM--TIF 207
QY 132 KEDGNILGHKLEVYNSHNYIMADKONGIKVNFKIRHNIEDGSVQLADHYQNTPIGD 191
Db 208 TASRSLGKEMRDLDTPFAFLYSDLDKGFPIINF--VEPNPLMEHYRRDHAQAIS--- 263
QY 192 GPVLLPDNNHISTQSALSKDPNEKEDHWLLEFVTAAGITLGMDELYKLSHGPPPEVEE 251
Db 264 -----GTWMSLKERERKND-----IQNRDL---IDELMKNSTYKDGTKMTD 302
QY 252 QDDGTLPWSCAQESGMDRHPAACA 277
Db 303 QETANLLIGVLMGC---QHTSAATSA 325

```

Search completed: May 14, 2004, 08:25:46
 Job time : 20 secs

Db 61 VTTFSYGVQCSRRYDHNKRDHFKFSAMPEGVQERTIFFPKDDGNKYTRAEVKEGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 180
 QY 162 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 239
 Db 181 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 238

RESULT 2

Q93125 PRELIMINARY; PRT; 238 AA.
 AC Q93125;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Green fluorescent protein mutant 3.
 GN GFP.
 OS Aequorea victoria (Jellyfish).
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequoridae; Aequorea.
 OX NCBI_TaxID=6100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96305137; PubMed=8707053;
 RA Cormack B.P., Valdivia R.H., Falkow S.;
 RT "FACS-optimized mutants of the green fluorescent protein (GFP).";
 RL Gene 173:33-38 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cormack B.P., Berttram G., Egerton M., Gow N.A.R., Falkow S.,
 RA Brown A.J.P.;
 RT "Yeast Enhanced Green Fluorescent Protein (YEGFP): a reporter of gene
 expression in Candida albicans.";
 RL Microbiology 0:0-0 (1996).
 DR EMBL; U73901; AAB18957.1; -.
 DR HSSP; P42312; 1BPP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPUDRSCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26840 MW; A28622809A9DEA6q CRC64;

Query Match 83.0%; Score 1245; DB 5; Length 238;
 Best Local Similarity 97.9%; Pred. No. 5.4e-95;

Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATTYGLTLKFICTTGKLPVPMPTL 61
 Db 1 MSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATTYGLTLKFICTTGKLPVPMPTL 60
 QY 62 VTTLYGVQCSRRYDHNKRDHFKFSAMPEGVQERTIFFPKDDGNKYTRAEVKEGDTLV 121
 Db 61 VTTFSYGVQCSRRYDHNKRDHFKFSAMPEGVQERTIFFPKDDGNKYTRAEVKEGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 180
 QY 162 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 239
 Db 181 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 238

RESULT 3

Q8GHE4 PRELIMINARY; PRT; 238 AA.
 AC Q8GHE4;
 DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Green fluorescence protein.
 GN 375GFP.
 OS Azomonas agilis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azomonas.
 OX NCBI_TaxID=116849;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koranyi P., Berenyi M., Burg K.;
 RT "Occurrence of green fluorescence protein in diazotrophic bacteria
 Azomonas and Azotobacter";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF324405; AAN86137.1; -.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPUDRSCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26902 MW; 15FE9B9C5B4F6B89 CRC64;

Query Match 82.8%; Score 1242; DB 2; Length 238;
 Best Local Similarity 97.5%; Pred. No. 9.5e-95;

Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATTYGLTLKFICTTGKLPVPMPTL 61
 Db 1 MSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATTYGLTLKFICTTGKLPVPMPTL 60
 QY 62 VTTLYGVQCSRRYDHNKRDHFKFSAMPEGVQERTIFFPKDDGNKYTRAEVKEGDTLV 121
 Db 61 VTTFSYGVQCSRRYDHNKRDHFKFSAMPEGVQERTIFFPKDDGNKYTRAEVKEGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEFYNNSHNVIYIMADKONGIKVNFKIRHNIEDGSVOLAD 180
 QY 162 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 239
 Db 181 HYQONTPIGDGPVLLPDNNHYLSTQSALSKDPNEKEDHVLLEFVTAAGITLGMDELYK 238

RESULT 4

Q8GHE3 PRELIMINARY; PRT; 238 AA.
 AC Q8GHE3;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Green fluorescence protein.
 GN 85GFP.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koranyi P., Berenyi M., Burg K.;
 RT "Occurrence of green fluorescence protein in diazotrophic bacteria
 Azomonas and Azotobacter";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF324406; AAN86138.1; -.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPUDRSCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26887 MW; B0E1616BD2AF6188 CRC64;

Query Match 82.7%; Score 1241; DB 2; Length 238;

Best Local Similarity 97.5%; Pred. No. 1,1e-94;
Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
DB 1 MSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 121
DB 61 VTTESYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 120
QY 122 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 181
DB 121 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

RESULT 5

Q17105 PRELIMINARY; PRT; 238 AA.

AC Q17105; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroidea; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RE SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X83959; CA58789.1; -.
DR PIR; J50692; J01514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPDRESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26950 MM; 26E2BB450E748B44 CRC64;

Query Match 80.6%; Score 1209; DB 5; Length 238;
Best Local Similarity 94.1%; Pred. No. 5.1e-92;
Matches 224; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
DB 1 MSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 121
DB 61 VTTESYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 120
QY 122 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 181
DB 121 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

RESULT 6
Q17106 PRELIMINARY; PRT; 238 AA.

AC Q17106;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroidea; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RE SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X83960; CA58790.1; -.
DR PIR; J50692; J01514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPDRESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26867 MM; BD4648262DBEABD4 CRC64;

Query Match 79.6%; Score 1194; DB 5; Length 238;
Best Local Similarity 93.3%; Pred. No. 8.9e-91;
Matches 222; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
DB 1 MSKGEELFTGVPLVFLVDGVDNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 121
DB 61 VTTESYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKTRAEVKEGGTIV 120
QY 122 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 181
DB 121 NRILKGIIDPEKDNIIIGHKLEYNNSHNVIYIMADKQNGIKVFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

RESULT 7

Q8WTC6 PRELIMINARY; PRT; 238 AA.

AC Q8WTC6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroidea; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RE SEQUENCE FROM N.A.
RA STRAIN=GFPxm19uv;
RA Liu W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435431; AL33916.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.

DR PRINTS; PRO1229; GFPUSCENT.
DR Prodom; PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 27015 MW; 688FD75B88926903 CRC64;

Query Match 72.6%; Score 1089; DB 5; Length 238;
Best Local Similarity 82.4%; Pred. No. 4.3e-82;
Matches 196; Conservative 21; Mismatches 21; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVQCFRSRYPEDHMKOHDFPKSAMPEGVQERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTSGYQCFARVYEHKMDFFKSAMPEGVIGERTIFFQDDGKTKRGVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEGGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 238

RESULT 8

Q8WP95 PRELIMINARY; PRT; 238 AA.

AC Q8WP95;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFPxm.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GFPxm, and GFPdnam;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
RA Li S.J., Xia N.S.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY013824; AAK02062.1; -
DR EMBL; AY013821; AAK02059.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP. 1.
DR PRINTS; PRO1229; GFPUSCENT.
DR Prodom; PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 27049 MW; 8185D0B5E529012B CRC64;

Query Match 72.3%; Score 1085; DB 5; Length 238;
Best Local Similarity 81.9%; Pred. No. 9.2e-82;
Matches 195; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVQCFRSRYPEDHMKOHDFPKSAMPEGVQERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTSGYQCFARVYEHKMDFFKSAMPEGVIGERTIFFQDDGKTKRGVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEGGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 238

RESULT 9

Q8WTC4 PRELIMINARY; PRT; 238 AA.

AC Q8WTC4;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shg24;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435433; AAL33918.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP. 1.
DR PRINTS; PRO1229; GFPUSCENT.
DR Prodom; PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 26956 MW; 75521EAF0CEBA73A CRC64;

Query Match 72.1%; Score 1082; DB 5; Length 238;
Best Local Similarity 83.2%; Pred. No. 1.6e-81;
Matches 198; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVQCFRSRYPEDHMKOHDFPKSAMPEGVQERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTSGYQCFARVYEHKMDFFKSAMPEGVIGERTIFFQDDGKTKRGVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKRHNIEGGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELTK 238

RESULT 10

Q8WTD0 PRELIMINARY; PRT; 238 AA.

AC Q8WTD0;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GFPxm; 61;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GFPxmi19uv;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla."
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435430; AAL33915.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 27002 MW; BDSBA2982264C018 CRC64;

Query Match 71.7%; Score 1076; DB 5; Length 238;
 Best Local Similarity 81.5%; Pred. No. 5.1e-81;
 Matches 194; Conservative 21; Mismatches 23; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNGHKSFSVSGEGDADYGLTKLFTCTGKLPVMPPTL 61
 DB 1 MSKGEELFTGIVPVLLIEDGDVHGKFSVRGSGEGDADYGLKIFICTTGKLPVMPPTL 60
 QY 62 VTTLTGVOCFSRYPDHMKOHDPFKSAMPEGYVQERTIFFKDDGNKTRAEVKEGDTLV 121
 DB 61 VTTLTGVLICFARYPEHMKMNDPFKSAMPEGYIQERTIFFQDDGXKKTGGEVKEGDTLV 120
 QY 122 NRLEKIGDFKEDNIIIGHKLEYNNSHNYIMADKQNGIKVNFIRNIEDSGVQIAD 181
 DB 121 NRLEKIGDFKEDNIIIGHKLEYNNSHNYIMADKQNGIKVNFIRNIEDSGVQIAD 180
 QY 182 HYQONTPIGDGVLLPDNHYLSTQSAISKDPNEKRDHWLLEFVYAGITLGMDELK 239
 DB 181 HYQNTVPLGDGVLLPINHYLSTQSAISKDPNEKRDHWLLEFVYAGITLGMDELK 238

RESULT 14
 Q8WTC5 PRELIMINARY; PRT; 238 AA.
 ID Q8WTC5;
 AC Q8WTC5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Orange fluorescent protein.
 GN GFP.
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequorea; Aequorea.
 NC NCB1_TaxID=147615;
 RX NCB1_TaxID=147615;
 RN SEQUENCE FROM N.A.
 RC STRAIN=OFPxm;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla."
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435432; AAL33917.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 27018 MW; 75521EAS534E573A CRC64;

Query Match 71.7%; Score 1075; DB 5; Length 238;
 Best Local Similarity 82.8%; Pred. No. 6.2e-81;
 Matches 197; Conservative 17; Mismatches 24; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNGHKSFSVSGEGDADYGLTKLFTCTGKLPVMPPTL 61

DB 1 MSKGEELFTGVVPIVLVDGVDNGHKSFSVSGEGDADYGLTKLFTCTGKLPVMPPTL 60
 QY 62 VTTLTGVOCFSRYPDHMKOHDPFKSAMPEGYVQERTIFFKDDGNKTRAEVKEGDTLV 121
 DB 61 VTTLTGVLICFARYPEHMKMNDPFKSAMPEGYIQERTIFFQDDGXKKTGGEVKEGDTLV 120
 QY 122 NRLEKIGDFKEDNIIIGHKLEYNNSHNYIMADKQNGIKVNFIRNIEDSGVQIAD 181
 DB 121 NRLEKIGDFKEDNIIIGHKLEYNNSHNYIMADKQNGIKVNFIRNIEDSGVQIAD 180
 QY 182 HYQONTPIGDGVLLPDNHYLSTQSAISKDPNEKRDHWLLEFVYAGITLGMDELK 239
 DB 181 HYQNTVPLGDGVLLPINHYLSTQSAISKDPNEKRDHWLLEFVYAGITLGMDELK 238

RESULT 15
 Q95UA7 PRELIMINARY; PRT; 225 AA.
 ID Q95UA7;
 AC Q95UA7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cyan fluorescent protein (Fragment).
 OS Monasteria cavernosa (great star coral).
 OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Scleractinia;
 OC Faviina; Faviidae; Monasteria.
 NC NCB1_TaxID=63558;
 RX NCB1_TaxID=63558;
 RN SEQUENCE FROM N.A.
 RA Falkowski P.G., Sun Y.;
 RT "Monasteria cavernosa fluorescent protein."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY056460; AAL17905.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 FT NON TER 225
 SQ SEQUENCE 225 AA; 25775 MW; 52DE2F716D083524 CRC64;

Query Match 16.8%; Score 251.5; DB 5; Length 225;
 Best Local Similarity 31.1%; Pred. No. 7.8e-13;
 Matches 64; Conservative 43; Mismatches 82; Indels 17; Gaps 7;

QY 12 VPIVLVDGVDNGHKSFSVSGEGDADYGLTKLFTCTGKLPVMPPTLVTTLTYGQ 70
 DB 7 VKIKLRDGIYNGHKSFSVSGEGDADYGLTKLFTCTGKLPVMPPTLVTTLTYGQ 66
 QY 71 CFSRYPDHMKOHDPFKSAMPEGYVQERTIFFKDDGNKTRAEVKEGDTLVNRLEKIG 130
 DB 67 VFTKPKDIP--DYRKQSFPEGSWSERSMTFEDQVCVCTSDIKLEGSCFFYEIRFYGVN 124
 QY 131 FKEDNIIIGHKLEYNNSHNYIMADKQNGIKVNFIRNIEDSGVQIADHYQONTPI 189
 DB 125 FSSGDPVQKTKLTKMPESTENNTV----RDGVLLGDVSRITLLRGD---KHKCNRRS 175
 QY 190 GDGP---VLLPDNHYLSTQ-SALSKD 211
 DB 176 TYGAKKGVLLPEYHFVDRIRIETLSHD 201

Search completed: May 14, 2004, 08:26:43
 Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:59 ; Search time 22 Seconds
(Without alignments)
659,404 Million cell updates/sec

Title: US-09-931-232-1
Sequence: 1 MWSKGELFLGVPIVLELD.....AQSGMDRHPACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCBUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500	100.0	281	US-09-062-102-1	Sequence 1, Appl1
2	1500	100.0	281	US-09-364-946-1	Sequence 1, Appl1
3	1289.5	86.0	805	US-09-513-783A-178	Sequence 178, App
4	1280.5	85.4	1171	US-09-417-197-131	Sequence 131, App
5	1279	85.3	997	US-09-417-197-121	Sequence 121, App
6	1278	85.2	607	US-09-417-197-115	Sequence 115, App
7	1276	85.1	544	US-09-417-197-115	Sequence 115, App
8	1276	85.1	631	US-09-417-197-39	Sequence 39, Appl
9	1276	85.1	843	US-09-417-197-117	Sequence 117, App
10	1275.5	85.0	933	US-09-417-197-135	Sequence 135, App
11	1275.5	85.0	941	US-09-513-783A-172	Sequence 172, App
12	1275	85.0	727	US-09-417-197-139	Sequence 139, App
13	1275	85.0	797	US-09-417-197-143	Sequence 143, App
14	1274	84.9	239	US-09-172-063-3	Sequence 3, Appl1
15	1274	84.9	239	US-09-513-783A-46	Sequence 46, Appl1
16	1274	84.9	239	US-09-316-919-4	Sequence 4, Appl1
17	1274	84.9	239	US-09-602-641-3	Sequence 2, Appl1
18	1274	84.9	239	US-09-920-922-2	Sequence 2, Appl1
19	1274	84.9	294	US-09-513-783A-2	Sequence 2, Appl1
20	1274	84.9	323	US-09-172-063-21	Sequence 21, Appl
21	1274	84.9	323	US-09-602-641-21	Sequence 21, Appl
22	1274	84.9	364	US-09-085-305-6	Sequence 6, Appl1
23	1274	84.9	379	US-09-417-197-129	Sequence 129, App
24	1274	84.9	432	US-09-800-170-48	Sequence 48, Appl
25	1274	84.9	459	US-09-417-197-127	Sequence 127, App
26	1274	84.9	459	US-09-513-783A-170	Sequence 170, App
27	1274	84.9	544	US-09-417-197-113	Sequence 113, App

28	1274	84.9	604	US-09-417-197-59	Sequence 59, Appl
29	1274	84.9	605	US-09-417-197-41	Sequence 41, Appl
30	1274	84.9	606	US-09-417-197-65	Sequence 65, Appl
31	1274	84.9	630	US-09-417-197-63	Sequence 63, Appl
32	1274	84.9	633	US-09-417-197-45	Sequence 45, Appl
33	1274	84.9	635	US-09-417-197-125	Sequence 125, App
34	1274	84.9	642	US-08-818-253-6	Sequence 2, Appl1
35	1274	84.9	642	US-08-818-253-2	Sequence 6, Appl1
36	1274	84.9	642	US-08-818-252-2	Sequence 2, Appl1
37	1274	84.9	642	US-08-818-252-6	Sequence 6, Appl1
38	1274	84.9	652	US-08-818-253-4	Sequence 4, Appl1
39	1274	84.9	652	US-08-818-252-4	Sequence 4, Appl1
40	1274	84.9	718	US-09-417-197-75	Sequence 75, Appl
41	1274	84.9	719	US-09-417-197-51	Sequence 51, Appl
42	1274	84.9	726	US-09-417-197-71	Sequence 71, Appl
43	1274	84.9	783	US-09-513-783A-176	Sequence 176, App
44	1274	84.9	797	US-09-417-197-141	Sequence 141, App
45	1274	84.9	798	US-09-417-197-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1
US-09-062-102-1
Sequence 1, Application US/09062102
Patent No. 6130313
GENERAL INFORMATION:
APPLICANT: Kain, Steve
APPLICANT: Li, Xiangjiang
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
FILE REFERENCE: D6100
CURRENT APPLICATION NUMBER: US/09/062,102
CURRENT FILING DATE: 1998-04-17
EARLIER APPLICATION NUMBER: US 60/060, 855
EARLIER FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SEQ ID NO 1
LENGTH: 281
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
Patent No. 6130313
US-09-062-102-1

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.2e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKGELFLGVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKICTTGKLPVPMPT 60
DB 1 MWSKGELFLGVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKICTTGKLPVPMPT 60
QY 61 LVTTLVYGVOCFSRYPDHNKOHDFKSNAMPEGVQRTIFFKDDGNYKTRAYPEGDTL 120
DB 61 LVTTLVYGVOCFSRYPDHNKOHDFKSNAMPEGVQRTIFFKDDGNYKTRAYPEGDTL 120
QY 121 VNIIEIKGIDFKEDGNITGKLEYNNSHNVYIMAKOKNGIKVNFIRINIDGSVQLA 180
DB 121 VNIIEIKGIDFKEDGNITGKLEYNNSHNVYIMAKOKNGIKVNFIRINIDGSVQLA 180
QY 181 DHYQONTPIGDGVLLPDNHYLSTQSLAKDPNEKRDHWLLEFVTAAGITLGMDELK 240
DB 181 DHYQONTPIGDGVLLPDNHYLSTQSLAKDPNEKRDHWLLEFVTAAGITLGMDELK 240
QY 241 LSHGFPPEVEEDGTIPMSCAQESGMDRHPACASARINV 281
DB 241 LSHGFPPEVEEDGTIPMSCAQESGMDRHPACASARINV 281

RESULT 2

US-09-364-946-1
 ; Sequence 1, Application US/09364946
 ; Patent No. 6306600
 ; GENERAL INFORMATION:
 ; APPLICANT: Kain, Steve
 ; APPLICANT: Li, Xiangqiang
 ; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
 ; TITLE OF INVENTION: of Use
 ; FILE REFERENCE: D6100CIP/D2
 ; CURRENT APPLICATION NUMBER: US/09/364,946
 ; EARLIER FILING DATE: 1999-07-30
 ; EARLIER APPLICATION NUMBER: US 09/191,233
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 1
 ; LENGTH: 281
 ; TYPE: PRT
 ; ORGANISM: artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
 ; Patent No. 6306600
 US-09-364-946-1

Query Match 100.0%; Score 1500; DB 4; Length 281;
 Best Local Similarity 100.0%; Pred. No. 1.2e-150;
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Cy 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60

Cy 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120
    |||
Db 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240

Cy 241 LSHGPPPEVEBODDGLTPMSCAQSBSGMDRHPAACAASARINV 281
    |||
Db 241 LSHGPPPEVEBODDGLTPMSCAQSBSGMDRHPAACAASARINV 281
  
```

RESULT 3
 US-09-513-783A-178
 ; Sequence 178, Application US/09513783A
 ; Patent No. 6416959
 ; GENERAL INFORMATION:
 ; APPLICANT: Giuliano, Kenneth A.
 ; APPLICANT: Kapur, Ravi
 ; TITLE OF INVENTION: A System for Cell Based Screening
 ; FILE REFERENCE: 97-022-11
 ; CURRENT APPLICATION NUMBER: US/09/513,783A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 178
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: GFP-NFKB
 ; Patent No. 6416959
 US-09-513-783A-178

Query Match 86.0%; Score 1289.5; DB 4; Length 805;
 Best Local Similarity 92.9%; Pred. No. 1.2e-127;
 Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

```

Cy 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60

Cy 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120
    |||
Db 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240

Cy 241 LSHGPPPEVEBODDGLTPMSCAQSBSG 266
    |||
Db 241 GLSRDPPFMEDELPLIFPAEPAQASG 267
  
```

RESULT 4
 US-09-417-197-131
 ; Sequence 131, Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP, et al.
 ; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
 ; TITLE OF INVENTION: On A Cellular Response
 ; FILE REFERENCE: 3759-0110P
 ; CURRENT APPLICATION NUMBER: US/09/417,197
 ; CURRENT FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 131
 ; LENGTH: 1171
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: EGFP-NFAT fusion
 ; Patent No. 6518021
 US-09-417-197-131

Query Match 85.4%; Score 1280.5; DB 4; Length 1171;
 Best Local Similarity 93.9%; Pred. No. 1.9e-126;
 Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;

```

Cy 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGBDATYGLTLTKFTCTTGKLPVPMPT 60

Cy 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120
    |||
Db 61 LVTLTYGVCCFSRYPDHMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAAYKFBGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGVPILPDNHYLSTOSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240

Cy 241 LSHGPPPEVEBODDGLTPMSCAQSBSG 255
    |||
Db 241 GLSRDPPFMEDELPLIFPAEPAQASG 258
  
```

RESULT 5
 US-09-417-197-121
 ; Sequence 121, Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP, et al.

```

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 121
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-lkappab-kinase fusion
US-09-417-197-121

```

```

Query Match      85.3%; Score 1279; DB 4; Length 997;
Best Local Similarity 90.4%; Pred. No. 2.2e-126;
Matches 245; Conservative 2; Mismatches 10; Indels 14; Gaps 1;

```

```

QY 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
DB 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120
DB 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120
QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK- 239
QY 241 LSHGFPEPEVEBDDGTLPMSCAQESGMDRHP 271
DB 240 -----GLRSRAQAQNSMTMERPP 257

```

RESULT 6

```

US-09-417-197-47
; Sequence 47, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 47
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-p38 fusion
US-09-417-197-47

```

```

Query Match      85.2%; Score 1278; DB 4; Length 607;
Best Local Similarity 84.7%; Pred. No. 1.3e-126;
Matches 249; Conservative 6; Mismatches 15; Indels 24; Gaps 3;

```

```

QY 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
DB 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120
DB 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120

```

```

QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK- 239
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKS 240
QY 240 -----KLSHGFPEPEVEBDDGTLPMSCAQESGMDRHPAACS 276
DB 241 GLRSRGAQNSMTMERPP-----GSGAYGSGVCA 288

```

RESULT 7

```

US-09-417-197-115
; Sequence 115, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 115
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-CDK2 fusion
US-09-417-197-115

```

```

Query Match      85.1%; Score 1276; DB 4; Length 544;
Best Local Similarity 93.5%; Pred. No. 1.8e-126;
Matches 244; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

```

```

QY 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
DB 1 MWSKGEELFTGVVILVLELDGVNKHKFSVSGEGGATYGLTKLTKFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120
DB 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAPEGYOERTIFFKDDGNKTRAVAFEGDYL 120
QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK- 239
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKS 240
QY 240 -----KLSHGFPEPEVEBDDGT 256
DB 241 GLRSRAQNSMTMERPP-----GSGAYGSGVCA 260

```

RESULT 8

```

US-09-417-197-39
; Sequence 39, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 39

```

LENGTH: 631
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-Etk1 fusion construct
US-09-417-197-39

Query Match 85.1%; Score 1276; DB 4; Length 631;
Best Local Similarity 91.7%; Pred. No. 2.2e-126;
Matches 244; Conservative 4; Mismatches 14; Indels 4; Gaps 2;

QY 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
QY 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
DB 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
QY 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
QY 241 LSHGPPPEVERODDGLTLMSCAQBSG 266
DB 241 --GHSRRA-QASNSTMAAAAGGG 262

RESULT 9
US-09-417-197-117
Sequence 117; Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole Thastrup, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 117
LENGTH: 843
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-PTP fusion
US-09-417-197-117

Query Match 85.1%; Score 1276; DB 4; Length 843;
Best Local Similarity 92.8%; Pred. No. 3.5e-126;
Matches 246; Conservative 2; Mismatches 9; Indels 8; Gaps 2;

QY 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
QY 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
DB 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
QY 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240

QY 241 -----LSHG-PPPEVERODDGLT 257
DB 241 GLASREMLSRGMFHRDLSDGLDATT 265

RESULT 10
US-09-417-197-135
Sequence 135; Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole Thastrup, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 135
LENGTH: 933
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-PK9 fusion
US-09-417-197-135

Query Match 85.0%; Score 1275.5; DB 4; Length 933;
Best Local Similarity 91.7%; Pred. No. 4.6e-126;
Matches 242; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

QY 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNKHKFSVSGEGDATTGKLTIKFICTTGKLPVWPT 60
QY 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
DB 61 LVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEBGT 120
QY 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPFKEKDNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
QY 241 -----LSHGPPPEVER 251
DB 241 GLSRGSMGTIRDLQYALQEXIEE 264

RESULT 11
US-09-513-783A-172
Sequence 172; Application US/09513783A
Patent No. 6416959
GENERAL INFORMATION:
APPLICANT: Giuliano, Kenneth A.
TITLE OF INVENTION: A System for Cell Based Screening
FILE REFERENCE: 97-022-111
CURRENT APPLICATION NUMBER: US/09/513,783A
CURRENT FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 180
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 172
LENGTH: 941
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GFP-HSP70
US-09-513-783A-172

Query Match 85.0%; Score 1275.5; DB 4; Length 941;

Best Local Similarity 85.0%; Pred. No. 4.6e-126;
Matches 250; Conservative 3; Mismatches 10; Indels 31; Gaps 4;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
Db 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
QY 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
Db 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
QY 241 -----LSHGFPPEVEEODDGTILPMGC-----AQSNG-----DRHPAACS 276
Db 241 GMSVGVGIDIGF-----QSCYAVARAGIETIANEYSDRCPACIS 281

RESULT 12

US-09-417-197-139
; Sequence 139, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin Version 3.0
; SEQ ID NO 139
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PK8 fusion
US-09-417-197-139

Query Match 85.0%; Score 1275; DB 4; Length 727;

Best Local Similarity 94.2%; Pred. No. 3.5e-126;
Matches 244; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
Db 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
QY 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
Db 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
QY 241 -----LSHGFPPEVEEODDGTIL 257
Db 241 GLRSRGTMDSVAIVKEGWL 259

RESULT 13

US-09-417-197-143
; Sequence 143, Application US/09417197

; Patent No. 6518021

; GENERAL INFORMATION:

; APPLICANT: Ole THASTRUP, et al.

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An

; FILE REFERENCE: 3759-0110P

; CURRENT APPLICATION NUMBER: US/09/417,197

; CURRENT FILING DATE: 1999-10-07

; NUMBER OF SEQ ID NOS: 143

; SOFTWARE: Patentin Version 3.0

; SEQ ID NO 143

; LENGTH: 797

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: EGFP-NFKappaB fusion
US-09-417-197-143

Query Match 85.0%; Score 1275; DB 4; Length 797;

Best Local Similarity 90.4%; Pred. No. 4.1e-126;
Matches 246; Conservative 1; Mismatches 11; Indels 14; Gaps 2;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGSDATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
Db 61 LVTTLVYGVQCFSRYPDHMKOHDFEKSAMPEGYVOERTTFPKDDGNYKTRAIVKFEEDTL 120
QY 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
Db 121 VNRLEKGDIFKEDGNILGHKLEYNNSHNVTIMADKOKNGIKVNFKIRHNIEDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSXPNEKRDHWLLEFVTAAGITLGMDELYKK 240
QY 241 -----LSHGFPPEVEEODDGTILPMSCAQESG 266
Db 241 GLRSRAMELIFPL-----IFPAEPQASG 264

RESULT 14

US-09-172-063-3

; Sequence 3, Application US/09172063

; Patent No. 6150176

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.

; APPLICANT: Miyawaki, Atsushi

; APPLICANT: Llopis, Juan

; APPLICANT: Machter, Rebekka M.

; APPLICANT: Remington, S. James

; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR

; FILE REFERENCE: 07257/071001

; CURRENT APPLICATION NUMBER: US/09/172,063

; CURRENT FILING DATE: 1998-10-13

; EARLIER APPLICATION NUMBER: 09/094,359

; EARLIER FILING DATE: 1998-06-09

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 239

; TYPE: PRT

; ORGANISM: Aequorea victoria

; FEATURE:
; NAME/KEY: VARIANT

; LOCATION: (0)...(0)

; OTHER INFORMATION: EGFP
US-09-172-063-3

Query Match 84.9%; Score 1274; DB 3; Length 239;

Best Local Similarity 100.0%; Pred. No. 8.2e-127;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
QY 61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNKTRAVEKFEEDTL 120
Db 61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNKTRAVEKFEEDTL 120
QY 121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNEFKIRHNIEDGSVOLA 180
Db 121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNEFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
Db 181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
```

RESULT 15

```
US-09-513-783A-46
; Sequence 46 Application US/09513783A
; Patent No. 6416859
; GENERAL INFORMATION:
; APPLICANT: Giuliani, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-11
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: EGFP
US-09-513-783A-46
```

Query Match 84.9%; Score 1274; DB 4; Length 239;

Best Local Similarity 100.0%; Pred. No. 8.2e-127;

Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
QY 61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNKTRAVEKFEEDTL 120
Db 61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGVQERTIFFKDDGNKTRAVEKFEEDTL 120
QY 121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNEFKIRHNIEDGSVOLA 180
Db 121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNEFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
Db 181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
```

Search completed: May 14, 2004, 08:27:50
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:26:49 ; Search time 570 Seconds
(Without alignments)
137.178 Million cell updates/sec

Title: US-09-931-232-1
Perfect score: 1500
Sequence: 1 MWSKGELFGVPIVLELD.....AQSMDRRHACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09C_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	100.0	281	US-09-931-232-1	Sequence 1, Appl1
2	1290.5	86.0	288	US-10-314-861-37	Sequence 37, Appl1
3	1290.5	86.0	293	US-10-314-861-35	Sequence 35, Appl1
4	1290.5	86.0	295	US-10-314-861-39	Sequence 39, Appl1
5	1290.5	86.0	299	US-10-314-861-33	Sequence 33, Appl1
6	1290.5	86.0	305	US-10-314-861-31	Sequence 31, Appl1
7	1290.5	86.0	311	US-10-314-861-29	Sequence 29, Appl1
8	1290.5	86.0	324	US-10-314-861-16	Sequence 16, Appl1
9	1289.5	86.0	805	US-10-100-957A-178	Sequence 178, App
10	1279	85.4	1171	US-10-072-036-131	Sequence 131, App
11	1279	85.3	997	US-10-072-036-121	Sequence 121, App
12	1278	85.2	1089	US-10-072-036-47	Sequence 47, Appl1
13	1277	85.1	1090	US-10-259-864-2	Sequence 2, Appl1
14	1276.5	85.1	1090	US-10-259-864-6	Sequence 6, Appl1
15	1276	85.1	544	US-10-072-036-115	Sequence 115, App

16	1276	85.1	631	US-10-072-036-39	Sequence 39, Appl1
17	1276	85.1	843	US-10-072-036-117	Sequence 117, App
18	1275.5	85.0	933	US-10-072-036-135	Sequence 135, App
19	1275.5	85.0	941	US-10-100-957A-172	Sequence 172, App
20	1275	85.0	727	US-10-072-036-139	Sequence 139, App
21	1275	85.0	797	US-10-072-036-143	Sequence 143, App
22	1274	84.9	239	US-09-920-922-2	Sequence 2, Appl1
23	1274	84.9	239	US-09-999-745-4	Sequence 4, Appl1
24	1274	84.9	239	US-09-866-538-4	Sequence 4, Appl1
25	1274	84.9	239	US-09-797-496B-2	Sequence 2, Appl1
26	1274	84.9	239	US-09-794-308-4	Sequence 4, Appl1
27	1274	84.9	239	US-09-865-291-4	Sequence 4, Appl1
28	1274	84.9	239	US-10-457-982-3	Sequence 3, Appl1
29	1274	84.9	239	US-10-121-258-13	Sequence 13, Appl1
30	1274	84.9	239	US-10-221-461-7	Sequence 7, Appl1
31	1274	84.9	239	US-10-100-957A-2	Sequence 2, Appl1
32	1274	84.9	239	US-10-177-390-2	Sequence 2, Appl1
33	1274	84.9	239	US-10-338-411-3	Sequence 3, Appl1
34	1274	84.9	239	US-10-370-570-4	Sequence 4, Appl1
35	1274	84.9	239	US-10-389-640-3	Sequence 3, Appl1
36	1274	84.9	259	US-10-314-861-11	Sequence 11, Appl1
37	1274	84.9	294	US-10-100-957A-2	Sequence 2, Appl1
38	1274	84.9	308	US-10-033-717-35	Sequence 35, Appl1
39	1274	84.9	320	US-10-338-411-11	Sequence 11, Appl1
40	1274	84.9	320	US-10-389-640-11	Sequence 11, Appl1
41	1274	84.9	323	US-10-457-982-21	Sequence 21, Appl1
42	1274	84.9	323	US-10-338-411-7	Sequence 7, Appl1
43	1274	84.9	323	US-10-338-411-13	Sequence 13, Appl1
44	1274	84.9	323	US-10-389-640-7	Sequence 7, Appl1
45	1274	84.9	323	US-10-389-640-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-931-232-1
Sequence 1, Application US/09931232
Publication No. US20020058274A1
GENERAL INFORMATION:
APPLICANT: Li, Xiangliang
APPLICANT: Kain, Steve
TITLE OR INVENTION: Rapidly Degrading GFP-Fusion Proteins
FILE REFERENCE: CLONING5CON
CURRENT APPLICATION NUMBER: US/09/931,232
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/364,946
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/191,233
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 09/062,102
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 60/060,855
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 281
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
US-09-931-232-1

Query Match 100.0%; Score 1500; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.7e-144;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MWSKGELFGVPIVLELDGVDNGHGFYSGBEGDGYGKLTIKFTCTTGKLPVWPPT 60
Db 1 MWSKGELFGVPIVLELDGVDNGHGFYSGBEGDGYGKLTIKFTCTTGKLPVWPPT 60

QY 61 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
QY 241 LSHGPPPEVEEODDGLTLMSCAQSSGMDRHPAACASARINV 281
DB 241 LSHGPPPEVEEODDGLTLMSCAQSSGMDRHPAACASARINV 281

RESULT 2
US-10-314-861-37
; Sequence 37, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-ER3:980-1008
US-10-314-861-37

Query Match 86.0%; Score 1290.5; DB 14; Length 288;
Best Local Similarity 95.7%; Pred. No. 4,4e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYGLTKFTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYGLTKFTCTTGKLPVPMPT 80
QY 61 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 120
DB 81 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 140
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEXKED 273

RESULT 3
US-10-314-861-35
; Sequence 35, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-ER3:980-1013
US-10-314-861-35

Query Match 86.0%; Score 1290.5; DB 14; Length 293;
Best Local Similarity 95.7%; Pred. No. 4,5e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYGLTKFTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYGLTKFTCTTGKLPVPMPT 80
QY 61 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 120
DB 81 LVTLTYGVQCFSPYDPMKOHDFPKSAMPEGYVOERTIFFKDDGNKYKTRAEVFEEDTL 140
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEKGIIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEXKED 273
RESULT 4
US-10-314-861-39
; Sequence 39, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: US/09/270,427
 PRIOR FILING DATE: 1999-03-16
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: FASTSEQ For Windows Version 4.0
 SEQ ID NO 39
 LENGTH: 295
 TYPE: PRT
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1008
 US-10-314-861-39

Query Match 86.0%; Score 1290.5; DB 14; Length 295;
 Best Local Similarity 95.7%; Pred. No. 4,6e-123;
 Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 60
 DB 21 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 80
 QY 61 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 120
 DB 81 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 140
 QY 121 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
 DB 141 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
 QY 181 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 DB 201 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKS 260
 QY 241 LSHGFPPEVEEODD 254
 DB 261 -GGAGAPRIEKKED 273

RESULT 5

US-10-314-861-33
 Sequence 33, Application US/10314861
 Publication No. US20030148269A1
 GENERAL INFORMATION:
 APPLICANT: Abbott Laboratories
 APPLICANT: Holzman, Thomas F.
 APPLICANT: Harlan, John Eric
 APPLICANT: Egan, David A.
 APPLICANT: Buko, Alexander M.
 APPLICANT: Solomon, Larry R.
 APPLICANT: Lador, Uri S.
 APPLICANT: Tang, Qiang
 TITLE OF INVENTION: CENTRIFUGALITY-ENHANCED METHOD OF
 TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
 FILE REFERENCE: 6404.US.O1
 CURRENT APPLICATION NUMBER: US/10/314,861
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: US/09/270,427
 PRIOR FILING DATE: 1999-03-16
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: FASTSEQ For Windows Version 4.0
 SEQ ID NO 33
 LENGTH: 299
 TYPE: PRT
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1019
 US-10-314-861-33

Query Match 86.0%; Score 1290.5; DB 14; Length 299;
 Best Local Similarity 95.7%; Pred. No. 4,6e-123;
 Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 60
 |||

DB 21 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 80
 QY 61 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 120
 DB 81 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 140
 QY 121 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
 DB 141 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
 QY 181 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 DB 201 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKS 260
 QY 241 LSHGFPPEVEEODD 254
 DB 261 -GGAGAPRIEKKED 273

RESULT 6

US-10-314-861-31
 Sequence 31, Application US/10314861
 Publication No. US20030148269A1
 GENERAL INFORMATION:
 APPLICANT: Abbott Laboratories
 APPLICANT: Holzman, Thomas F.
 APPLICANT: Harlan, John Eric
 APPLICANT: Egan, David A.
 APPLICANT: Buko, Alexander M.
 APPLICANT: Solomon, Larry R.
 APPLICANT: Lador, Uri S.
 APPLICANT: Tang, Qiang
 TITLE OF INVENTION: CENTRIFUGALITY-ENHANCED METHOD OF
 TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
 FILE REFERENCE: 6404.US.O1
 CURRENT APPLICATION NUMBER: US/10/314,861
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: US/09/270,427
 PRIOR FILING DATE: 1999-03-16
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: FASTSEQ For Windows Version 4.0
 SEQ ID NO 31
 LENGTH: 305
 TYPE: PRT
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1025
 US-10-314-861-31

Query Match 86.0%; Score 1290.5; DB 14; Length 305;
 Best Local Similarity 95.7%; Pred. No. 4,6e-123;
 Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 60
 DB 21 MWSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGDGYATYGLTKLFTCTTGKLPVWPPT 80
 QY 61 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 120
 DB 81 LVTLITGVQCFSRYPDHMKQDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTL 140
 QY 121 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
 DB 141 VNRTELKIDFKEDGNILGHLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
 QY 181 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKK 240
 DB 201 DHYOQNTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKS 260
 QY 241 LSHGFPPEVEEODD 254
 DB 261 -GGAGAPRIEKKED 273

```
RESULT 7
US-10-314-861-29
; Sequence 29, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.O1
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1031
US-10-314-861-29
```

```
Query Match      86.0%; Score 1290.5; DB 14; Length 311;
Best Local Similarity 95.7%; Pred. No. 4,9e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 60
DB 21 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 80
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 120
DB 81 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 140
QY 121 VNRLEKGDIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEKGDIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYK 240
DB 201 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYK 260
QY 241 LSHGFPPEVEODD 254
DB 261 -GGAGGRIRKED 273
```

```
RESULT 8
US-10-314-861-16
; Sequence 16, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.O1
; CURRENT APPLICATION NUMBER: US/10/314,861
```

```
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1044
US-10-314-861-16
```

```
Query Match      86.0%; Score 1290.5; DB 14; Length 324;
Best Local Similarity 95.7%; Pred. No. 5,2e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 60
DB 21 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 80
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 120
DB 81 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 140
QY 121 VNRLEKGDIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEKGDIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYK 240
DB 201 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYK 260
QY 241 LSHGFPPEVEODD 254
DB 261 -GGAGGRIRKED 273
```

```
RESULT 9
US-10-100-957A-178
; Sequence 178, Application US/10100957A
; Publication No. US20030096322A1
; GENERAL INFORMATION:
; APPLICANT: Giuliani, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-11A
; CURRENT APPLICATION NUMBER: US/10/100,957A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GFP-NFKB
US-10-100-957A-178
```

```
Query Match      86.0%; Score 1289.5; DB 14; Length 805;
Best Local Similarity 92.9%; Pred. No. 2,5e-122;
Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

QY 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 60
DB 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDAATYGLTKLTKCTTGKLPVPWPT 60
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 120
DB 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVOERTIFFKDDGNYKTRAEVKFEGDTL 120
QY 121 VNRLEKGDIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
```

Db 121 VNRIELKIDFKEDGNILGHKLENYNSHNVIYIMADKQKNGIKVNFKIRHINIEDGSVOJA 180
 QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 Db 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 QY 241 -LSHGFPPEVEODDGTLPMSCAOESG 266
 Db 241 GLRSRDPPEWDELFLIFPAAPQASG 267

RESULT 10

US-10-072-036-131
 ; Sequence 131, Application US/10072036
 ; Publication No. US20030082564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP
 ; APPLICANT: Sara BJORN
 ; APPLICANT: Soren TULLIN
 ; APPLICANT: Kasper ALMHOLT
 ; APPLICANT: Kurt SCUDDER
 ; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
 ; FILE REFERENCE: 3759-0120P
 ; CURRENT APPLICATION NUMBER: US/10/072,036
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 09/417,197
 ; PRIOR FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 131
 ; LENGTH: 1171
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: EGFP-NFAT fusion
 US-10-072-036-131

Query Match 85.4%; Score 1280.5; DB 14; Length 1171;
 Best Local Similarity 93.9%; Pred. No. 3.6e-121;
 Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;

QY 1 MVSKEELFTGVVPIIVELDDGVNGHKFVSGBEGSDATYKLTLEICTTGKLPVMPPT 60
 Db 1 MVSKEELFTGVVPIIVELDDGVNGHKFVSGBEGSDATYKLTLEICTTGKLPVMPPT 60
 QY 61 LVTTLYGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKXDGNYKTRAEVKFGDTL 120
 Db 61 LVTTLYGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKXDGNYKTRAEVKFGDTL 120
 QY 121 VNRIELKIDFKEDGNILGHKLENYNSHNVIYIMADKQKNGIKVNFKIRHINIEDGSVOJA 180
 Db 121 VNRIELKIDFKEDGNILGHKLENYNSHNVIYIMADKQKNGIKVNFKIRHINIEDGSVOJA 180
 QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 Db 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 QY 241 LSHGFPPEVEODDGTLPMSCAOESG 266
 Db 241 LSHGFPPEVEODDGTLPMSCAOESG 266
 QY 241 LSHGFPPEVEODDGTLPMSCAOESG 266
 Db 241 LSHGFPPEVEODDGTLPMSCAOESG 266

RESULT 11

US-10-072-036-121
 ; Sequence 121, Application US/10072036
 ; Publication No. US20030082564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP
 ; APPLICANT: Sara BJORN
 ; APPLICANT: Soren TULLIN
 ; APPLICANT: Kasper ALMHOLT
 ; APPLICANT: Kurt SCUDDER

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An In
 ; FILE REFERENCE: 3759-0120P
 ; CURRENT APPLICATION NUMBER: US/10/072,036
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 09/417,197
 ; PRIOR FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 121
 ; LENGTH: 997
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: EGFP-lkappab-kinase fusion
 US-10-072-036-121

Query Match 85.3%; Score 1279; DB 14; Length 997;
 Best Local Similarity 90.4%; Pred. No. 4.1e-121;
 Matches 245; Conservative 2; Mismatches 10; Indels 14; Gaps 1;

QY 1 MVSKEELFTGVVPIIVELDDGVNGHKFVSGBEGSDATYKLTLEICTTGKLPVMPPT 60
 Db 1 MVSKEELFTGVVPIIVELDDGVNGHKFVSGBEGSDATYKLTLEICTTGKLPVMPPT 60
 QY 61 LVTTLYGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKXDGNYKTRAEVKFGDTL 120
 Db 61 LVTTLYGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKXDGNYKTRAEVKFGDTL 120
 QY 121 VNRIELKIDFKEDGNILGHKLENYNSHNVIYIMADKQKNGIKVNFKIRHINIEDGSVOJA 180
 Db 121 VNRIELKIDFKEDGNILGHKLENYNSHNVIYIMADKQKNGIKVNFKIRHINIEDGSVOJA 180
 QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 Db 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKDHVLLFEVTAAGITLGMDELYK 240
 QY 241 LSHGFPPEVEODDGTLPMSCAOESGMDRHP 271
 Db 241 LSHGFPPEVEODDGTLPMSCAOESGMDRHP 271

RESULT 12

US-10-072-036-47
 ; Sequence 47, Application US/10072036
 ; Publication No. US20030082564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP
 ; APPLICANT: Sara BJORN
 ; APPLICANT: Soren TULLIN
 ; APPLICANT: Kasper ALMHOLT
 ; APPLICANT: Kurt SCUDDER
 ; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An In
 ; FILE REFERENCE: 3759-0120P
 ; CURRENT APPLICATION NUMBER: US/10/072,036
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 09/417,197
 ; PRIOR FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 47
 ; LENGTH: 607
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: EGFP-p38 fusion
 US-10-072-036-47

Query Match 85.2%; Score 1278; DB 14; Length 607;
 Best Local Similarity 84.7%; Pred. No. 2.5e-121;
 Matches 249; Conservative 6; Mismatches 15; Indels 24; Gaps 3;

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
QY 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX- 239
DB 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 240 -----LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAS 276
DB 241 GHSRGRKMSQERPTFYROELNKTIWEVERYQNLSPV-----GSGAYGSVCAA 288
```

```
RESULT 13
US-10-259-864-2
; Sequence 2, Application US/10259864
; Publication No. US2003007645A1
; GENERAL INFORMATION:
; APPLICANT: Hager, Gordon L
; TITLE OF INVENTION: Superfamily Receptor Chimeras, Translocation Assay For Superfamily
; TITLE OF INVENTION: Receptor Ligands, and Methods and Kits For Detecting and Character
; TITLE OF INVENTION: Ligands
; FILE REFERENCE: 24799
; CURRENT APPLICATION NUMBER: US/10/259,864
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/325,178
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Rat/human chimera
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..()
; OTHER INFORMATION: Chimeric Protein
US-10-259-864-2
```

```
Query Match 85.1%; Score 1277; DB 14; Length 1089;
Best Local Similarity 86.4%; Pred. No. 7,4e-121;
Matches 247; Conservative 8; Mismatches 25; Indels 6; Gaps 2;
```

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
QY 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
DB 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 241 -----LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 GALLDSKESLAPPGRDEVPGL-LGGGRGSVMDFYSLRGATVKV 285
```

RESULT 14

```
US-10-259-864-6
; Sequence 6, Application US/10259864
; Publication No. US2003007645A1
; GENERAL INFORMATION:
; APPLICANT: Hager, Gordon L
; TITLE OF INVENTION: Superfamily Receptor Chimeras, Translocation Assay For Superfamily
; TITLE OF INVENTION: Receptor Ligands, and Methods and Kits For Detecting and Character
; TITLE OF INVENTION: Ligands
; FILE REFERENCE: 24799
; CURRENT APPLICATION NUMBER: US/10/259,864
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/325,178
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1090
; TYPE: PRT
; ORGANISM: Rat/human Chimera
US-10-259-864-6
```

```
Query Match 85.1%; Score 1276.5; DB 14; Length 1090;
Best Local Similarity 86.1%; Pred. No. 8,4e-121;
Matches 247; Conservative 8; Mismatches 25; Indels 7; Gaps 2;
```

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYOEERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNHYIMADKQNGIKVNFKIRHNIEDGSYOLA 180
QY 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
DB 181 DHYQONTPIGDGPVLLPNDHNYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 241 -----LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 GALLDSKESLAPPGRDEVPGL-LGGGRGSVMDFYSLRGATVKV 286
```

```
RESULT 15
US-10-072-036-115
; Sequence 115, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An It
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 115
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-CDK2 fusion
US-10-072-036-115
```

Query Match 85.1%; Score 1276; DB 14; Length 544;
 Best Local Similarity 93.5%; Pred. No. 3.4e-121;
 Matches 244; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

QY	1	MVSKGEELFTGVVPILVELDGDVNGHKFVS	60
Db	1	MVSKGEELFTGVVPILVELDGDVNGHKFVS	60
QY	61	LVTTLTGVQCFSRYPDMKQHDFFKSAMPE	120
Db	61	LVTTLTGVQCFSRYPDMKQHDFFKSAMPE	120
QY	121	VNRTELKGIIDFKEDGNILGHKLEYNNSH	180
Db	121	VNRTELKGIIDFKEDGNILGHKLEYNNSH	180
QY	181	DHYQONTPIGDPVLLPDNHYLSTQSALS	239
Db	181	DHYQONTPIGDPVLLPDNHYLSTQSALS	239
QY	240	----KLHGPPPEVEEDDGT	256
Db	241	GLRSRAMENF-QKVEKIGEGT	260

Search completed: May 14, 2004, 08:41:20
 Job time : 570 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:25:19 ; Search time 183 Seconds

(Without alignments)
1498.743 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500

Sequence: 1 MVSKEHLEFGVPIVLELD.....AQSGMDRHPACASARINV 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapevt 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/prodata/2/paa/US06_PCTUS_COMB.pep.*
2: /cgn2_6/prodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/prodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/prodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/prodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/prodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/prodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/prodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/prodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/prodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/prodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/prodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/prodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/prodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/prodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/prodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/prodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/prodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/prodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/prodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/prodata/2/paa/US097A_COMB.pep.*
22: /cgn2_6/prodata/2/paa/US097B_COMB.pep.*
23: /cgn2_6/prodata/2/paa/US098_COMB.pep.*
24: /cgn2_6/prodata/2/paa/US099A_COMB.pep.*
25: /cgn2_6/prodata/2/paa/US099B_COMB.pep.*
26: /cgn2_6/prodata/2/paa/US100_COMB.pep.*
27: /cgn2_6/prodata/2/paa/US101_COMB.pep.*
28: /cgn2_6/prodata/2/paa/US102_COMB.pep.*
29: /cgn2_6/prodata/2/paa/US103_COMB.pep.*
30: /cgn2_6/prodata/2/paa/US104_COMB.pep.*
31: /cgn2_6/prodata/2/paa/US106_COMB.pep.*
32: /cgn2_6/prodata/2/paa/US107_COMB.pep.*
33: /cgn2_6/prodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

1	1500	100.0	281	1	PCT-US98-24323-1	Sequence 1, Appl1
2	1500	100.0	281	15	US-09-191-233-1	Sequence 1, Appl1
3	1500	100.0	281	15	US-09-191-233-1	Sequence 1, Appl1
4	1500	100.0	281	17	US-09-365-088-1	Sequence 1, Appl1
5	1500	100.0	281	24	US-09-931-232-1	Sequence 1, Appl1
6	1290.5	86.0	288	16	US-09-270-427A-37	Sequence 37, Appl1
7	1290.5	86.0	288	29	US-09-270-427A-35	Sequence 35, Appl1
8	1290.5	86.0	293	16	US-09-270-427A-35	Sequence 35, Appl1
9	1290.5	86.0	293	29	US-10-314-861-35	Sequence 35, Appl1
10	1290.5	86.0	295	16	US-09-270-427A-39	Sequence 39, Appl1
11	1290.5	86.0	295	29	US-10-314-861-39	Sequence 39, Appl1
12	1290.5	86.0	299	16	US-09-270-427A-33	Sequence 33, Appl1
13	1290.5	86.0	299	29	US-10-314-861-33	Sequence 33, Appl1
14	1290.5	86.0	305	16	US-09-270-427A-31	Sequence 31, Appl1
15	1290.5	86.0	305	29	US-10-314-861-31	Sequence 31, Appl1
16	1290.5	86.0	311	16	US-09-270-427A-29	Sequence 29, Appl1
17	1290.5	86.0	311	29	US-10-314-861-29	Sequence 29, Appl1
18	1290.5	86.0	324	16	US-09-270-427A-16	Sequence 16, Appl1
19	1290.5	86.0	324	29	US-10-314-861-16	Sequence 16, Appl1
20	1289.5	86.0	805	27	US-10-100-957A-178	Sequence 178, Appl1
21	1280.5	85.4	1171	26	US-10-072-036-131	Sequence 131, Appl1
22	1279.5	85.3	997	26	US-10-072-036-121	Sequence 121, Appl1
23	1278	85.2	607	26	US-10-072-036-47	Sequence 47, Appl1
24	1277	85.1	1089	1	PCT-US02-31043-2	Sequence 2, Appl1
25	1276.5	85.1	1089	28	US-10-259-864-2	Sequence 6, Appl1
26	1276.5	85.1	1090	1	PCT-US02-31043-6	Sequence 6, Appl1
27	1276.5	85.1	1090	28	US-10-259-864-6	Sequence 6, Appl1
28	1276.5	85.1	1090	28	US-10-259-864-6	Sequence 6, Appl1
29	1276	85.1	289	1	PCT-US03-03550-70	Sequence 68, Appl1
30	1276	85.1	290	1	PCT-US03-03550-72	Sequence 70, Appl1
31	1276	85.1	290	1	PCT-US03-03550-74	Sequence 72, Appl1
32	1276	85.1	544	26	US-10-072-036-115	Sequence 74, Appl1
33	1276	85.1	631	23	US-09-807-345-6	Sequence 115, Appl1
34	1276	85.1	631	26	US-10-072-036-39	Sequence 39, Appl1
35	1276	85.1	843	26	US-10-072-036-117	Sequence 117, Appl1
36	1275.5	85.0	943	26	US-60-499-955-56	Sequence 56, Appl1
37	1275.5	85.0	943	26	US-10-072-036-135	Sequence 135, Appl1
38	1275.5	85.0	941	27	US-10-100-957A-172	Sequence 172, Appl1
39	1275	85.0	432	1	PCT-US02-13092-25	Sequence 25, Appl1
40	1275	85.0	727	26	US-10-072-036-139	Sequence 139, Appl1
41	1275	85.0	727	23	US-09-807-345-14	Sequence 14, Appl1
42	1275	85.0	727	26	US-10-072-036-143	Sequence 143, Appl1
43	1275	85.0	918	23	US-09-807-345-18	Sequence 18, Appl1
44	1274.5	85.0	255	27	US-10-129-207A-8	Sequence 8, Appl1
45	1274	84.9	229	1	PCT-US00-13684-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1

PCT-US98-24323-1

Sequence 1, Application PC/TUS9824323A

GENERAL INFORMATION:

APPLICANT: Kain, Steve

APPLICANT: Li, Xiangqiang

TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods

TITLE OF INVENTION: of use

FILE REFERENCE: D6100PCT

CURRENT APPLICATION NUMBER: PCT/US98/24323A

CURRENT FILING DATE: 1998-11-13

EARLIER APPLICATION NUMBER: US 09/062,102

EARLIER FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 3

SEQ ID NO 1

LENGTH: 281

TYPE: PRT

ORGANISM: artificial sequence

FEATURE:

OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.

PCT-US98-24323-1

Query Match 100.0%; Score 1500; DB 1; Length 281;

Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
QY 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
DB 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
```

RESULT 2

US-09-191-233-1
; Sequence 1, Application US/09191233A
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP
; CURRENT APPLICATION NUMBER: US/09/191,233A
; EARLIER FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: US 09/062,102
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-191-233-1

Query Match 100.0%; Score 1500; DB 15; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
QY 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
DB 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
```

RESULT 3
US-09-191-233-1

; Sequence 1, Application US/09191233B

; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP
; CURRENT APPLICATION NUMBER: US/09/191,233B
; EARLIER FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: US 09/062,102
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-191-233-1

Query Match 100.0%; Score 1500; DB 15; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
QY 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
DB 61 LVTTLTLYGQCFSRYPDHMKQHDFFKSAMPEGVYQERTIFFDDGNGYKTRAEVKEGDTL 120
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEEODDGTLPMSCAQESGMDRHPAACASARINV 281
```

RESULT 4
US-09-365-089-1

; Sequence 1, Application US/09365089
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP/D1
; CURRENT APPLICATION NUMBER: US/09/365,089
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/191,233
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-365-089-1

Query Match 100.0%; Score 1500; DB 17; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTLLKFTCTTGKLPVPMPT 60
```



```

Db      1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
QY      61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
        61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
Db      121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
        121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
Db      121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
        121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
QY      181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
        181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db      181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
        181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
QY      241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281
        241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281

```

RESULT 5

```

US-09-931-232-1
; Sequence 1, Application US/09931232
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangiang
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: C10075CON
; CURRENT APPLICATION NUMBER: US/09/931,232
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/364,946
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/191,233
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 09/062,102
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 60/060,855
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
; OTHER INFORMATION: protein.
US-09-931-232-1

```

```

Query Match      100.0%; Score 1500; DB 24; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
        1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
Db      1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
        1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
QY      61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
        61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
Db      61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
        61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
QY      121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
        121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
Db      121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
        121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
QY      181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
        181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db      181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
        181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
QY      241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281
        241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281

```

RESULT 6

```

US-09-270-427A-37
; Sequence 37, Application US/09270427A
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiong
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/09/270,427A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-09-270-427A-37

```

```

Query Match      86.0%; Score 1290.5; DB 16; Length 288;
Best Local Similarity 95.7%; Pred. No. 3.6e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

```

```

QY      1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
        1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
Db      21 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 80
        21 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 80
QY      61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
        61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 120
Db      81 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 140
        81 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKYKTRAEVKKEGDTL 140
QY      121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
        121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 180
Db      141 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 200
        141 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGSVOIA 200
QY      181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
        181 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db      201 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKS 260
        201 DRYQONTPIGDGPVLLPDNHYLSTQSALSQDNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY      241 LSHGFPPEVEEODD 254
        241 LSHGFPPEVEEODD 254
Db      261 -GGAGPRIKKED 273
        261 -GGAGPRIKKED 273

```

RESULT 7

```

US-10-314-861-37
; Sequence 37, Application US/10314861
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiong
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39

```

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 288
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-10-314-861-37

Query Match 86.0%; Score 1290.5; DB 29; Length 288;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 80
QY 61 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 120
DB 81 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 140
QY 121 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

RESULT 8

US-09-270-427A-35
Sequence 35, Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404.US.01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 293
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1013
US-09-270-427A-35

Query Match 86.0%; Score 1290.5; DB 16; Length 293;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 80
QY 61 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 120
DB 81 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 140

QY 121 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

RESULT 9

US-10-314-861-35
Sequence 35, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404.US.01
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US/09/270,427
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 293
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1013
US-10-314-861-35

Query Match 86.0%; Score 1290.5; DB 29; Length 293;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTKLKEICTTGKLPVPMPT 80
QY 61 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 120
DB 81 LVTLTYGVQCFSRYPDHMKQHDFFKSAWPEGYQERTIFFKDDGNKTAAYKFEEDTL 140
QY 121 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260

QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

RESULT 10

US-09-270-427A-39
Sequence 39, Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.

APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLING OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 295
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1008
US-09-270-427A-39

Query Match 86.0%; Score 1290.5; DB 16; Length 295;
Best Local Similarity 95.7%; Pred. No. 3,7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 60
DB 21 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 80
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 120
DB 81 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 140
QY 121 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 141 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGGRIRLEKED 273

RESULT 11
US-10-314-861-39
Sequence 39, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLING OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US/09/270,427
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 295
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1008

US-10-314-861-39
Query Match 86.0%; Score 1290.5; DB 29; Length 295;
Best Local Similarity 95.7%; Pred. No. 3,7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 60
DB 21 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 80
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 120
DB 81 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 140
QY 121 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 141 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGGRIRLEKED 273

RESULT 12
US-09-270-427A-33
Sequence 33, Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLING OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1019
US-09-270-427A-33

Query Match 86.0%; Score 1290.5; DB 16; Length 299;
Best Local Similarity 95.7%; Pred. No. 3,8e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 60
DB 21 MVSKEELFTGVPIIVLELDGVNFKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 80
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 120
DB 81 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTL 140
QY 121 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
DB 141 VNRLELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260

QY 241 LSHGPPPEVEEODD 254
Db 261 -GGAGAPRIEXKED 273

RESULT 13
US-10-314-861-33

Sequence 33, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang

TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
FILE REFERENCE: 6404 US 01
CURRENT FILING DATE: 2002-12-09
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 33
LENGTH: 299
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1019
US-10-314-861-33

Query Match 86.0%; Score 1290.5; DB 29; Length 299;
Best Local Similarity 95.7%; Pred. No. 3.8e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNKHRSVSGEGDGYTKLTKFTCTTGKLPVMPPT 60
Db 21 MWSKGEELFTGVVPLIVELDGVNKHRSVSGEGDGYTKLTKFTCTTGKLPVMPPT 80
QY 61 LVTLITLVGVQCFSRYPDHMKQHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKEGDTL 120
Db 81 LVTLITLVGVQCFSRYPDHMKQHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKEGDTL 140
QY 121 VNRLEKIDPFKEGNIHLGKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSVQLA 180
Db 141 VNRLEKIDPFKEGNIHLGKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSVQLA 200
QY 181 DHYQONTPIGDGPVILPDNHYLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
Db 201 DHYQONTPIGDGPVILPDNHYLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
Db 261 -GGAGAPRIEXKED 273

RESULT 14
US-09-270-427A-31

Sequence 31, Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF

TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 31
LENGTH: 305
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1025
US-09-270-427A-31

Query Match 86.0%; Score 1290.5; DB 16; Length 305;
Best Local Similarity 95.7%; Pred. No. 3.9e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNKHRSVSGEGDGYTKLTKFTCTTGKLPVMPPT 60
Db 21 MWSKGEELFTGVVPLIVELDGVNKHRSVSGEGDGYTKLTKFTCTTGKLPVMPPT 80
QY 61 LVTLITLVGVQCFSRYPDHMKQHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKEGDTL 120
Db 81 LVTLITLVGVQCFSRYPDHMKQHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKEGDTL 140
QY 121 VNRLEKIDPFKEGNIHLGKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSVQLA 180
Db 141 VNRLEKIDPFKEGNIHLGKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSVQLA 200
QY 181 DHYQONTPIGDGPVILPDNHYLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKK 240
Db 201 DHYQONTPIGDGPVILPDNHYLSTQSALSKDPNKRKDHVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
Db 261 -GGAGAPRIEXKED 273

RESULT 15
US-10-314-861-31

Sequence 31, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 31
LENGTH: 305
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1025
US-10-314-861-31

Query Match 86.0%; Score 1290.5; DB 29; Length 305;
Best Local Similarity 95.7%; Pred. No. 3.9e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNKHRSVSGEGDGYTKLTKFTCTTGKLPVMPPT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:25:49 ; Search time 20 Seconds
(Without alignments)
492.121 Million cell updates/sec

Title: US-09-931-232-1

Sequence: 1 MVSKEBELFTGVPIIVLELD.....AQESGMDRHPACASARINV 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 236798 seqs, 35026417 residues

Total number of hits satisfying chosen parameters: 236798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1274.5	85.0	255	6	US-10-129-207B-8 Sequence 8, Appl1
2	1274	84.9	239	1	PCT-US04-10982-3 Sequence 3, Appl1
3	1274	84.9	386	6	US-10-811-081-97 Sequence 97, Appl1
4	1274	84.9	567	6	US-10-488-758-4 Sequence 4, Appl1
5	1260	84.0	491	1	PCT-US04-07434-58 Sequence 58, Appl1
6	1257	83.8	1100	1	PCT-US04-07434-48 Sequence 48, Appl1
7	1257	83.8	1132	1	PCT-US04-07434-46 Sequence 46, Appl1
8	1255	83.7	239	1	PCT-US04-07434-44 Sequence 44, Appl1
9	1255	83.4	1961	1	PCT-US04-07434-60 Sequence 60, Appl1
10	1251	83.4	238	5	US-09-786-377A-8 Sequence 8, Appl1
11	1251	83.4	238	6	US-10-333-680-3 Sequence 3, Appl1
12	1250	83.3	239	1	PCT-US04-10982-4 Sequence 4, Appl1
13	1250	83.3	239	1	PCT-US04-10982-8 Sequence 8, Appl1
14	1247	83.1	238	1	PCT-US04-10982-2 Sequence 2, Appl1
15	1247	83.1	238	6	US-10-491-013-2 Sequence 2, Appl1
16	1245	83.0	239	6	US-10-433-640-11 Sequence 11, Appl1
17	1245	83.0	238	1	PCT-US04-10982-5 Sequence 5, Appl1
18	1242	82.8	238	6	US-10-676-428-35 Sequence 35, Appl1
19	1242	82.8	238	6	US-10-491-013-3 Sequence 3, Appl1
20	1241	82.7	238	6	US-10-471-857-1 Sequence 1, Appl1
21	1240	82.7	239	1	PCT-US04-10982-6 Sequence 6, Appl1
22	1238	82.5	242	6	US-10-676-428-25 Sequence 25, Appl1
23	1238	82.5	239	6	US-10-676-428-1 Sequence 1, Appl1
24	1234	82.2	238	6	PCT-US04-10982-7 Sequence 7, Appl1
25	1233	82.2	238	6	US-10-676-428-26 Sequence 26, Appl1
26	1233	82.2	387	6	US-10-811-081-87 Sequence 87, Appl1

27	1232	82.1	238	6	US-10-676-428-29 Sequence 29, Appl1
28	1229	81.9	238	6	US-10-676-428-27 Sequence 27, Appl1
29	1228	81.9	238	6	US-10-676-428-28 Sequence 28, Appl1
30	1228	81.9	238	6	US-10-471-857-3 Sequence 3, Appl1
31	1227	81.8	238	6	US-10-676-428-36 Sequence 36, Appl1
32	1227	81.8	238	6	US-10-471-857-7 Sequence 7, Appl1
33	1224	81.6	238	6	US-10-676-428-23 Sequence 23, Appl1
34	1224	81.6	238	6	US-10-471-857-5 Sequence 5, Appl1
35	1223	81.5	238	6	US-10-676-428-31 Sequence 31, Appl1
36	1223	81.5	238	6	US-10-471-857-9 Sequence 9, Appl1
37	1219	81.3	238	6	US-10-676-428-32 Sequence 32, Appl1
38	1215	81.0	238	6	US-10-676-428-24 Sequence 24, Appl1
39	1214	80.9	238	6	US-10-676-428-30 Sequence 30, Appl1
40	1214	80.9	238	6	US-10-676-428-33 Sequence 33, Appl1
41	1214	80.9	238	6	US-10-471-857-11 Sequence 11, Appl1
42	1213	80.9	238	6	US-10-676-428-22 Sequence 22, Appl1
43	1212	80.8	238	6	US-10-676-428-37 Sequence 37, Appl1
44	1211	80.7	238	6	US-10-676-428-34 Sequence 34, Appl1
45	1210	80.7	238	6	US-10-471-857-13 Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-10-129-207B-8
; Sequence 8, Application US/10129207B
; GENERAL INFORMATION:
; APPLICANT: Boehr, Gerald
; APPLICANT: Esser, Dirk
; APPLICANT: Schmidt, Ulrich
; APPLICANT: ACGR Progenomics AG
; TITLE OF INVENTION: Method for Directed Packaging of Molecular Substances
; TITLE OF INVENTION: in Protein Shells
; FILE REFERENCE: 080377-000000US
; CURRENT APPLICATION NUMBER: US/10/129,207B
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: DE 199 52 982.5
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: WO PCT/EP00/10878
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: enhanced green
; OTHER INFORMATION: fluorescent protein (GFP) with fused C-terminal
; OTHER INFORMATION: proline-rich sequence (GFP-PRIP)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (246)..(255)
; OTHER INFORMATION: C-terminal proline-rich sequence
US-10-129-207B-8

Query Match 85.0%; Score 1274.5; DB 6; Length 255;

Best Local Similarity 97.2%; Pred. No. 4.3e-88;

Matches 241; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY	1	MVSKEBELFTGVPIIVLELDGVNKHFSVSGEGSDATYTKTLTKCTTGKLPVWPPT	60
DB	1	MVSKEBELFTGVPIIVLELDGVNKHGKFSVSGEGSDATYTKTLTKCTTGKLPVWPPT	60
QY	61	LVTTLYGVQCSTRYVDHKKQDPFSSAPPEGVQERTIFFDDGNYKRAVYKEGDTL	120
DB	61	LVTTLYGVQCSTRYVDHKKQDPFSSAPPEGVQERTIFFDDGNYKRAVYKEGDTL	120
QY	121	VRIELKGIQDFEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNRKIHNIEDGSVOLA	180
DB	121	VRIELKGIQDFEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNRKIHNIEDGSVOLA	180

QY 181 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 239
DB 181 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 240
QY 240 KLSHGPP 247
DB 241 RRASGPP 248

RESULT 2

PCT-US04-10982-3
Sequence 3, Application PC/TUS0410982

GENERAL INFORMATION:
APPLICANT: Odyssey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdi, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
TITLE OF INVENTION: ASSAYS
FILE REFERENCE: ODD1007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: Patent version 3.0
SEQ ID NO 3
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: "EGFP", an Aequorea-derived mutated variant of wild-type green fl
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al
TITLE: Creating new fluorescent probes for cell biology
JOURNAL: Nat. Rev. Mol. Cell Biol.
VOLUME: 3
ISSUE: 12
PAGES: 906-918
DATE: 2002-
RELEVANT RESIDUES: (1)..(239)
PCT-US04-10982-3

Query Match 84.9%; Score 1274; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 4.4e-88;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKEELFTGVVPLVLELDGPNVNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
DB 1 MWSKEELFTGVVPLVLELDGPNVNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
QY 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKDDGNKTRAEVKEGDTL 120
DB 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKDDGNKTRAEVKEGDTL 120
QY 121 VNRLEKIDPFKEKDNILGHLENNYNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPFKEKDNILGHLENNYNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 239
DB 181 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 239

RESULT 3

US-10-811-081-97
Sequence 97, Application US/10811081

GENERAL INFORMATION:
APPLICANT: El Teyar, Nebil
APPLICANT: Campbell, Robert K
APPLICANT: Kelton, Christie A
APPLICANT: He, Chaomei

TITLE OF INVENTION: Novel Glycoproteins and Methods of Use Thereof
FILE REFERENCE: 20993-003
CURRENT APPLICATION NUMBER: US/10/811,081
CURRENT FILING DATE: 2004-03-25
PRIOR APPLICATION NUMBER: US/10/457,047
PRIOR FILING DATE: 2003-06-05
PRIOR APPLICATION NUMBER: US/10/360,149
PRIOR FILING DATE: 2003-02-06
PRIOR APPLICATION NUMBER: US/09/927,876
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/225,035
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/202,724
PRIOR FILING DATE: 2000-05-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 97
LENGTH: 386
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-811-081-97

Query Match 84.9%; Score 1274; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 7.4e-88;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKEELFTGVVPLVLELDGPNVNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
DB 148 MWSKEELFTGVVPLVLELDGPNVNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 207
QY 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKDDGNKTRAEVKEGDTL 120
DB 208 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGYOERTIFFKDDGNKTRAEVKEGDTL 267
QY 121 VNRLEKIDPFKEKDNILGHLENNYNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
DB 268 VNRLEKIDPFKEKDNILGHLENNYNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 327

QY 181 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 239
DB 328 DHYQONTPIGDPVLLPNDHNYLSTOSALSKDNEKRDHMLLEFVTAAGITLGMDELYK 386

RESULT 4
US-10-488-758-4
Sequence 4, Application US/10488758

GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: Sailer, Andreas W.
APPLICANT: Pan, Jie
APPLICANT: Fong, Tung M.
APPLICANT: Marsh, Donald J.
TITLE OF INVENTION: MELANIN-CONCENTRATING HORMONE RECEPTOR
TITLE OF INVENTION: ANTAGONIST BINDING PROTEIN
FILE REFERENCE: 20782P
CURRENT APPLICATION NUMBER: US/10/488,758
CURRENT FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: PCT/US02/29931
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/325,129
PRIOR FILING DATE: 2001-09-26
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 567
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: MCH-1R antagonist binding protein
US-10-488-758-4

Query Match 84.9%; Score 1274; DB 6; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.1e-87;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVSKEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 60
DB 329 MVSKEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 388
QY 61 LVTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 120
DB 389 LVTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 448
QY 121 VNRLEKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
DB 449 VNRLEKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 508
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 239
DB 509 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 567

RESULT 5

PCT-US04-07434-58
Sequence 58, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 58
LENGTH: 491
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-58

Query Match 84.0%; Score 1260; DB 1; Length 491;
Best Local Similarity 89.0%; Pred. No. 1.1e-86;
Matches 242; Conservative 4; Mismatches 20; Indels 6; Gaps 2;

QY 2 VSKGEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 61
DB 1 MSKGEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 60
QY 62 VTTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 121
DB 61 VTTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 120
QY 122 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 181
DB 121 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 241
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 240
QY 242 SHGFP-----PEVEEKDDGTLPMS-CAQESGM 267
DB 241 NLRIPMASALPEALCSGSRTPVLCILLESV 272

RESULT 6
PCT-US04-07434-48

Sequence 48, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 48
LENGTH: 1100
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Rattus norvegicus
PCT-US04-07434-48

Query Match 83.8%; Score 1257; DB 1; Length 1100;
Best Local Similarity 97.5%; Pred. No. 4.2e-86;
Matches 235; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 61
DB 1 MSKGEELFTGVVPLIVELDDVNGHKFSVSGEGEGATYKLTIKFTCTTGKLPVMPPTL 60
QY 62 VTTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 121
DB 61 VTTLTYGVQCFSRYPDHMKOHDFPKSAMPEGVYQERTIPFKDGNKYTRAIVEFGDTL 120
QY 122 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 181
DB 121 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 241
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRBDHMLLEFVTAAGITLGMDELYK 240
QY 242 S 242
DB 241 A 241

RESULT 7
PCT-US04-07434-46

Sequence 46, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 46
LENGTH: 1132
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-46

Query Match 83.8%; Score 1257; DB 1; Length 1132;
Best Local Similarity 97.5%; Pred. No. 4.3e-86;
Matches 235; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 60
QY 62 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 121
DB 61 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 120
QY 122 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 241
DB 181 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 240
QY 242 S 242
DB 241 A 241

RESULT 8
PCT-US04-07434-44
; Sequence 44, Application PC/TUS0407434
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of California, The

; APPLICANT: Leonard, Rome H.
; APPLICANT: Valerie, Kickhoefer A.
; APPLICANT: Sujna, Raval-Fernandes
; APPLICANT: Phoebe, Stewart L.
; TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
; FILE REFERENCE: 14399-1PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07434
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 60/453,800
; PRIOR FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: completely synthesized
PCT-US04-07434-44

Query Match 83.7%; Score 1255; DB 1; Length 239;
Best Local Similarity 98.7%; Pred. No. 1.2e-86;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 60
QY 62 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 121
DB 61 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 120
QY 122 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 239
DB 181 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 238

RESULT 9
PCT-US04-07434-60

; Sequence 60, Application PC/TUS0407434
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of California, The
; APPLICANT: Leonard, Rome H.
; APPLICANT: Valerie, Kickhoefer A.
; APPLICANT: Sujna, Raval-Fernandes
; APPLICANT: Phoebe, Stewart L.
; TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
; FILE REFERENCE: 14399-1PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07434
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 60/453,800
; PRIOR FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
; LENGTH: 1961
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-60

Query Match 83.7%; Score 1255; DB 1; Length 1961;
Best Local Similarity 98.7%; Pred. No. 1.1e-85;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGSDATYKLTLPCTTGKLPVWPTL 60
QY 62 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 121
DB 61 VTTLYGVQCFSRYPDHMKOHDFKSNAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 120
QY 122 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNYIMADKXNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 239
DB 181 HYQONTPIGDPVLLPDMHYLSTOSALS KDPNEKRDMVLLFEVTAAGITLGMDELTK 238

RESULT 10
US-09-786-377A-8

; Sequence 8, Application US/09786377A
; GENERAL INFORMATION:
; APPLICANT: Loma Linda University
; APPLICANT: Szalay, Aladar A.
; APPLICANT: Wang, Yubao
; APPLICANT: Gefu, Wang-Pruski
; TITLE OF INVENTION: Method for Studying Protein Interactions in Vivo
; FILE REFERENCE: 11785-3
; CURRENT APPLICATION NUMBER: US/09/786,377A
; CURRENT FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: US 60/135,835
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: US 60/099,068
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: PCT/US99/20207
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: humanized green fluorescence protein cDNA
US-09-786-377A-8

Query Match 83.4%; Score 1251; DB 5; Length 238;

Best Local Similarity 98.3%; Pred. No. 2.3e-86;
Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPTL 60
QY 62 VTTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 121
DB 61 VTTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 120
QY 122 NRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQAD 181
DB 121 NRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 239
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 238

RESULT 11
US-10-333-680-3

Sequence 3, Application US/10333680
GENERAL INFORMATION:
APPLICANT: Geiger, Benjamin
APPLICANT: Kam, Zvi
APPLICANT: Zamlir, Eli
APPLICANT: Barshelsky, Alexander D.
APPLICANT: Shulman, Michael
APPLICANT: Ben-zeev, Avri
TITLE OF INVENTION: METHODS OF ISOLATING GENES ENCODING PROTEINS OF SPECIFIC
FILE REFERENCE: 25254
CURRENT APPLICATION NUMBER: US/10/333,680
CURRENT FILING DATE: 2003-10-06
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-10-333-680-3

Query Match 83.4%; Score 1251; DB 6; Length 238;
Best Local Similarity 98.3%; Pred. No. 2.3e-86;
Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPTL 60
QY 62 VTTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 121
DB 61 VTTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 120
QY 122 NRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQAD 181
DB 121 NRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 239
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 238

RESULT 12
PCT-US04-10982-4

Sequence 4, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odysey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdin, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT

TITLE OF INVENTION: ASSAYS
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial

OTHER INFORMATION: "EYFP", enhanced yellow fluorescent protein, an Aequorea-derived
OTHER INFORMATION: mutated variant of wild-type green fluorescent protein
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al
TITLE: Creating new fluorescent probes for cell biology
JOURNAL: Nat. Rev. Mol. Cell Biol.
VOLUME: 3
ISSUE: 3
PAGES: 906-918
DATE: 2002-
RELEVANT RESIDUES: (1) .. (239)
PCT-US04-10982-4

Query Match 83.3%; Score 1250; DB 1; Length 239;
Best Local Similarity 97.9%; Pred. No. 2.7e-86;
Matches 234; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPT 60
DB 1 MSKGEELFTGVVPLVLVDGVNKHKFSVSGEGDATTYKLTIKFTCTTGKLPVWPPT 60
QY 61 LVTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 120
DB 61 LVTTLYGVQCFSRYPDMKQHDFFKSAPEGYOERTIFFKDDGNKYKRAVKEFGDTLV 120
QY 121 VNRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQOLA 180
DB 121 VNRILKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSYQOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 239
DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVVLEFVTAAGITLGMDELK 239

RESULT 13
PCT-US04-10982-8

Sequence 8, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odysey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdin, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: "EGFP", cyan fluorescent protein, an Aequorea-derived mutated var
OTHER INFORMATION: tant of wt green fluorescent protei
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al

TITLE: Creating new fluorescent probes for cell biology
 JOURNAL: Nat. Rev. Mol. Cell Biol.
 VOLUME: 3
 ISSUE: 12
 PAGES: 906-918
 DATE: 2002-
 RELEVANT RESIDUES: (1)..(239)
 PCT-US04-10982-8

Query Match 83.1%; Score 1247; DB 1; Length 239;
 Best Local Similarity 97.9%; Pred. No. 4.5e-86;
 Matches 233; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSKEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 60
 DB 1 MVSKEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 60
 QY 61 LVTLLTVGQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 120
 DB 61 LVTLLTVGQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 120
 QY 121 VNRLEKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 DB 121 VNRLEKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
 DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239

RESULT 14
 PCT-US04-10982-2
 Sequence 2, Application PC/TUS0410982

GENERAL INFORMATION:
 APPLICANT: Odyssey Thera, Inc.
 APPLICANT: Michalek, Stephen
 APPLICANT: Macdonald, Marile
 APPLICANT: Lameridn, Jane
 TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
 TITLE OF INVENTION: ASSAYS
 FILE REFERENCE: ODDY007
 CURRENT APPLICATION NUMBER: PCT/US04/10982
 CURRENT FILING DATE: 2004-04-20
 PRIOR APPLICATION NUMBER: US 60/461,133
 PRIOR FILING DATE: 2003-04-09
 NUMBER OF SEQ ID NOS: 1067
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 2
 LENGTH: 238
 TYPE: PRT
 ORGANISM: Aequorea victoria
 PCT-US04-10982-2

Query Match 83.1%; Score 1247; DB 1; Length 238;
 Best Local Similarity 97.9%; Pred. No. 4.5e-86;
 Matches 233; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 61
 DB 1 MVSKEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 60
 QY 62 VTTLLTVGQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 121
 DB 61 VTTFSYVQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 120
 QY 122 NRIELKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 181
 DB 121 NRIELKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
 DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 238

RESULT 15
 US-10-491-013-2
 Sequence 2, Application US/10491013

GENERAL INFORMATION:
 APPLICANT: AMERSHAM BIOSCIENCES UK LIMITED
 APPLICANT: STUBBS, Simon
 APPLICANT: JONES, Anne B.
 APPLICANT: DAVIES, June
 APPLICANT: RUEHLIMANN, Dietrich
 APPLICANT: MICHAEL, Nigel P.
 TITLE OF INVENTION: FLUORESCENT PROTEINS
 FILE REFERENCE: PA0163
 CURRENT APPLICATION NUMBER: US/10/491,013
 CURRENT FILING DATE: 2004-03-25
 PRIOR APPLICATION NUMBER: GB 0123314.7
 PRIOR FILING DATE: 2001-09-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 2
 LENGTH: 238
 TYPE: PRT
 ORGANISM: Aequorea victoria
 US-10-491-013-2

Query Match 83.1%; Score 1247; DB 6; Length 238;
 Best Local Similarity 97.9%; Pred. No. 4.5e-86;
 Matches 233; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 61
 DB 1 MVSKEELFTGVPLIVELDDVNGHKFSVSGEGDGYGKLTLLKFTCTTGKLPVMPPTL 60
 QY 62 VTTLLTVGQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 121
 DB 61 VTTFSYVQCFSRYPDMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAEVFEEDTL 120
 QY 122 NRIELKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 181
 DB 121 NRIELKIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
 DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 238

Search completed: May 14, 2004, 08:31:38
 Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 23:59:10 ; Search time 3699 Seconds

(without alignments)
9959.875 Million cell updates/sec

Title: US-09-931-232-2

Sequence: 1 atggtgagcagcagcgagga.....tagatcaatgcttagatgc 850

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	850	100.0	850	6 AR112746	AR112746 Sequence
2	850	100.0	850	6 AR173740	AR173740 Sequence
3	850	100.0	850	6 BD232356	BD232356 Rapidly d
4	850	100.0	4280	6 AX923429	AX923429 Sequence
5	845	99.4	845	6 AX348042	AX348042 Sequence
6	845	99.4	845	6 AX353909	AX353909 Sequence
7	721.2	84.8	10417	6 AX617021	AX617021 Sequence
8	720.6	84.8	6418	6 AX490700	AX490700 Sequence
9	720.6	84.8	6418	6 AX801202	AX801202 Sequence
10	720.6	84.8	6418	6 AX814838	AX814838 Sequence
11	720	84.7	12789	6 AX590202	AX590202 Sequence
12	719.4	84.6	7326	6 AX814843	AX814843 Sequence
13	719.4	84.6	9952	6 AX801203	AX801203 Sequence
14	719.4	84.6	9952	6 AX814839	AX814839 Sequence
15	719.4	84.6	9952	6 AX814839	AX814839 Sequence
16	719.2	84.6	5713	6 BD222121	BD222121 Method an
17	719.2	84.6	5713	6 AX146811	AX146811 Sequence
18	719.2	84.6	8531	6 AX490699	AX490699 Sequence
19	719	84.6	8531	6 AX663051	AX663051 Sequence
20	719	84.6	8531	6 AX107919	AX107919 Sequence
21	719	84.6	10292	6 AX107920	AX107920 Sequence
22	719	84.6	10292	6 AX107917	AX107917 Sequence
23	719	84.6	10384	6 AX107918	AX107918 Sequence
24	719	84.6	10384	6 AX107918	AX107918 Sequence
25	719	84.6	10384	6 AX283601	AX283601 Sequence
26	718.8	84.6	6084	12 GGA575208	AJ575208 Expressio
27	718.6	84.5	1190	6 AX455471	AX455471 Sequence
28	718.6	84.5	1377	6 AX230579	AX230579 Sequence
29	718.6	84.5	1377	6 AX536308	AX536308 Sequence
30	718.6	84.5	3355	12 CVU76561	U76561 Cloning vec
31	718.6	84.5	4151	12 CVU55761	U55761 Cloning vec
32	718.6	84.5	4729	12 U57609	U57609 Cloning vec
33	718.6	84.5	4733	6 AX658018	AX658018 Sequence
34	718.6	84.5	4733	12 CVU55762	U55762 Cloning vec
35	718.6	84.5	4737	12 U57608	U57608 Cloning vec
36	718.6	84.5	4810	12 AF525449	AF525449 Expressio
37	718.6	84.5	4843	12 AF525447	AF525447 Expressio
38	718.6	84.5	4843	12 AF525448	AF525448 Expressio
39	718.6	84.5	5063	6 BD1339247	BD1339247 Method of
40	718.6	84.5	5184	6 BD270592	BD270592 Method fo
41	718.6	84.5	5451	6 AX771237	AX771237 Sequence
42	718.6	84.5	5497	12 AY237157	AY237157 EGFP expr
43	718.6	84.5	5619	6 AX008865	AX008865 Sequence
44	718.6	84.5	5619	6 BD211260	BD211260 Phagocyt
45	718.6	84.5	5710	12 AF490378	AF490378 Synthetic

ALIGNMENTS

RESULT 1
LOCUS AR112746 850 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6130313.
ACCESSION AR112746
VERSION AR112746.1 GI:14092646
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 850)
AUTHORS Li,X. and Kain,S.
TITLE Rapidly degrading GFP-fusion proteins
JOURNAL Patent: US 6130313-A 2 10-OCR-2000;
FEATURES Location/Qualifiers

QY	121	GGCAGAGTGAACCTCGAAGTTTCATCTGCAACACCGGAAAGCTGACCGTGCCTCGGCGCAAC	180
Db	217	GGCAGAGTGAACCTCGAAGTTTCATCTGCAACACCGGAAAGCTGACCGTGCCTCGGCGCAAC	276
QY	181	CTCGTGAACCACTCGACCTTAACGAGGTGACAGTGTTCAGCCCGCTAACCCGACCACTGAAG	240
Db	277	CTCGTGAACCACTCGACCTTAACGAGGTGACAGTGTTCAGCCCGCTAACCCGACCACTGAAG	336
QY	241	CAGCAACGACTTCTTCAGAGTCCGCCATGCCCCGAAAGCTACGTCAGAGAGCGACCATCTTC	300
Db	337	CAGCAACGACTTCTTCAGAGTCCGCCATGCCCCGAAAGCTACGTCAGAGAGCGACCATCTTC	396
QY	301	TTCAAGAGCAGCGGAACTTACAAAGACCCGCGCCGAGGTGAAGTTGAAGGGGAGCAACCCCTG	360
Db	397	TTCAAGAGCAGCGGAACTTACAAAGACCCGCGCCGAGGTGAAGTTGAAGGGGAGCAACCCCTG	456
QY	361	GTGAACCGCATCGAGCTGAAGAGGAGATCGACTTCAAGAGAGAGCGCAACATCTGGGGAC	420
Db	457	GTGAACCGCATCGAGCTGAAGAGGAGATCGACTTCAAGAGAGAGCGCAACATCTGGGGAC	516
QY	421	AAGCTGAGATCAACTACAAACGCCACAAGCTCTATATCATGSCCGACAAAGCAAGAAAC	480
Db	517	AAGCTGAGATCAACTACAAACGCCACAAGCTCTATATCATGSCCGACAAAGCAAGAAAC	576
QY	481	GGCATCAAGGTGAACCTCAAGATCCGACCAACATCGAGAGCGGAGGTGAGCTGCGC	540
Db	577	GGCATCAAGGTGAACCTCAAGATCCGACCAACATCGAGAGCGGAGGTGAGCTGCGC	636
QY	541	GACCACTAACGAGAAACACCCCATCGGCGACCGGCCCTGCTGCTCCGACCAACAC	600
Db	637	GACCACTAACGAGAAACACCCCATCGGCGACCGGCCCTGCTGCTCCGACCAACAC	696
QY	601	TACCTGAGACACCCAGTCCGCTTGAGCAAAAGACCCCAACAGAGAGCGGATACATGCTC	660
Db	697	TACCTGAGACACCCAGTCCGCTTGAGCAAAAGACCCCAACAGAGAGCGGATACATGCTC	756
QY	661	CTGCTGAGATTGAGAACCGCGCGGGAGTCACTCTGCGCATGGAAGAGCTGACAAAGAG	720
Db	757	CTGCTGAGATTGAGAACCGCGCGGGAGTCACTCTGCGCATGGAAGAGCTGACAAAGAG	816
QY	721	CTTAGACATGAGCTTCCCGCCGAGAGGTGAGAGACAGATGATGACACGCTGCCCATGCTC	780
Db	817	CTTAGACATGAGCTTCCCGCCGAGAGGTGAGAGACAGATGATGATGACACGCTGCCCATGCTC	876
QY	781	TGTGCCCCAGAGAGCGGGATGGAACCGTCAACCTGTGAAGCTTGTCTTCTGCTAGATCAAT	840
Db	877	TGTGCCCCAGAGAGCGGGATGGAACCGTCAACCTGTGAAGCTTGTCTTCTGCTAGATCAAT	936
QY	841	GTGTAGATGC 850	
Db	937	GTGTAGATGC 946	

RESULT 5	AX348042	845 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX348042				
DEFINITION	Sequence 21 from Patent Epi1172375.				
ACCESSION	AX348042				
VERSION	AX348042.1	GI:18614152			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
Source					

```

AX348042      845 bp      DNA      linear      PAT 06-FEB-2002
Sequence 21 from Patent Epi1172375.
AX348042
AX348042.1    GI:18614152
.
synthetic construct
synthetic construct
artificial sequences.
1
Odenthal, M. and Jung, D.
Gene expression, genome alteration and reporter gene expression in
myofibroblasts and myofibroblasts-like cells by using the
regulatory regions within the alpha smooth muscle alpha-actin gene
Patent: EP 1172375-A-21 16-JAN-2002;
Odenthal, Margarete (DE)
Location/Qualifiers
1..845
/organism="synthetic construct"

```

ORIGIN

```
/mol_type="unassigned DNA"  
/db_xref="taxon:32630"  
/note="EGFP"
```

Query Match	99.4%;	Score 845;	DB 6;	Length 845;
Best Local Similarity	100.0%;	Pred. No. 3.3e-98;		
Matches 845;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGCTGAGCAAGGGCGAGAGCTGTTCAACCGGGGTGGGCCATCTGTTGAGCTGGAC	60
Db	1	ATGCTGAGCAAGGGCGAGAGCTGTTCAACCGGGGTGGGCCATCTGTTGAGCTGGAC	60
QY	61	GGCGACGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCATGCACTTAC	120
Db	61	GGCGACGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCATGCACTTAC	120
QY	121	GGCAAGCTGACCCCTGAAGTTCACTGTCACACCGGCAAGCTGGCCCTGTGCGCCCAAC	180
Db	121	GGCAAGCTGACCCCTGAAGTTCACTGTCACACCGGCAAGCTGGCCCTGTGCGCCCAAC	180
QY	181	CTCGTGACCAACCTTGACCTTACGGGGTGAGTCTTCAAGCCGCTACCCCGACCACTGAAG	240
Db	181	CTCGTGACCAACCTTGACCTTACGGGGTGAGTCTTCAAGCCGCTACCCCGACCACTGAAG	240
QY	241	CAGAGACACTTCTTCAAGTCCGGCATGCGGAAGGCTACGTGTCAGAGAGCGGACATCTTC	300
Db	241	CAGAGACACTTCTTCAAGTCCGGCATGCGGAAGGCTACGTGTCAGAGAGCGGACATCTTC	300
QY	301	TTCAAGACGACGGGCAACTACAGACCCGCGCGAGGTGAAGTTTCAGAGGCGACACCTTG	360
Db	301	TTCAAGACGACGGGCAACTACAGACCCGCGCGAGGTGAAGTTTCAGAGGCGACACCTTG	360
QY	361	GTGAACCGCATCGAGCTGAGAGGGCATGACTTCAAGAGAGACGGCAACATCTGGGGAC	420
Db	361	GTGAACCGCATCGAGCTGAGAGGGCATGACTTCAAGAGAGACGGCAACATCTGGGGAC	420
QY	421	AAGCTGGAGTCAACTCTCAACAGGCAACAAGCTCTATATCATGGGCGGACAGAGCAAGAAC	480
Db	421	AAGCTGGAGTCAACTCTCAACAGGCAACAAGCTCTATATCATGGGCGGACAGAGCAAGAAC	480
QY	481	GGCATCAAGGTGAACCTTCAAGATCCGCCACCAACTGAGAGCGGACGGTGACCTCGCC	540
Db	481	GGCATCAAGGTGAACCTTCAAGATCCGCCACCAACTGAGAGCGGACGGTGACCTCGCC	540
QY	541	GACCATACACGACGAGACACCCCATGAGGAGACGGGCCCGGTGCTGTCCCGACCAACAC	600
Db	541	GACCATACACGACGAGACACCCCATGAGGAGACGGGCCCGGTGCTGTCCCGACCAACAC	600
QY	601	TACCTGAGCACCCAGTCCGCCCTGAGCAACACCCCAAGAGCGGATCACTGCTC	660
Db	601	TACCTGAGCACCCAGTCCGCCCTGAGCAACACCCCAAGAGCGGATCACTGCTC	660
QY	661	CTGCTGAGATTCTGTCACCGCCCGGGATCACTCTCCGATGAGACGACTGTACAAAGAG	720
Db	661	CTGCTGAGATTCTGTCACCGCCCGGGATCACTCTCCGATGAGACGACTGTACAAAGAG	720
QY	721	CTTAGACATGACTTCCGCGCGGAGGTGAGAGAGAGATGATGAGACGCTGCCATGTCT	780
Db	721	CTTAGACATGACTTCCGCGCGGAGGTGAGAGAGAGATGATGAGACGCTGCCATGTCT	780
QY	781	TGTGCCACGAGAGCGGGATGACCTGTCACTCTGACGCTTGCTCTTGCTAGATCAAT	840
Db	781	TGTGCCACGAGAGCGGGATGACCTGTCACTCTGACGCTTGCTCTTGCTAGATCAAT	840
QY	841	GTGTA 845	
Db	841	GTGTA 845	

RESULT 6	
AX353909	
LOCUS	
AX353909	845 bp
	DNA
	linear
	PAT 06-FEB-2002

[illegible]

RESULT 8			
AX490700			
LOCUS	AX490700	6418 bp	DNA
DEFINITION	Sequence 2 from Patent WO0229065.		linear
ACCESSION	AX490700		
VERSION	AX490700.1	GI:22323676	

SOURCE	ORGANISM
synthetic construct	synthetic construct
synthetic construct	artificial sequences.

AUTHORS
Kingsman, A. J., Mazarakis, N. D., Martin-Rendon, E., Azzouz, M. and Rohlf, J.

TITLE Retroviral vectors containing internal ribosomal entry sites
JOURNAL Patent: WO 0229065-A 2 11-APR-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES	LOCATION/Qualifiers
SOURCE	1. .6418

ORIGIN

Query Match:	84.8%	Score	720.6;	DB	6;	Length	6418;
Best Local Similarity:	96.8%	Pred	No.1.3e-82;				
Matches	735;	Conservative	0;	Mismatches	24;	Indels	0;
						Gaps	0

QY 1 ATGTGAGCAAGGGGAGGAGCTGTTACCGGGGGGTGCCATCTGTGAGCTGGAC 60

Db 2690 ATGTGAGCAAGGGGAGGAGCTGTTACCGGGGGGTGCCATCTGTGAGCTGGAC 2740

61 GGCGCAGTAAACGGCGACAAATTTCAGCTGTCGGCGAGGCGAGGCGCATGCCACTTAC 120

121 GGCAAGCTGACCCCTGAAGTTCATCTGCACCAACGGCAAGCTGCCCCGTGCCCTGGCCACC 180

CTGTGACCAACCCTGAAGCGTGCAGTGCTTCAGCCGGTACCCCAGACCATATGAAG

181 240

241 CAGCAGCACTTCTTCAAGTCCGCCATGCCGGAAGGCTTACGTCACAGGAGCGGACCATCTTC 300

301 TTCTAAGACGACGCAACTCAAGACCCCCGCGGAGTGAAGTTGAGGGGACACCTTG 360
Db 2930 CAGACGACTTCTTCAAGTCGGCCATGCCCAAGGCTACGTCAGGAGCGGACCACTTC 298

Db 2990 TTCAAGGACGACGGCAACTCAAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTG 304

QY	361	GTAAACCGCATCGAGCTGAAAGGGCATTCGATTCGAAGAGGACGGCAATCTCTGGGGCAC	420
Db	3050	GTAAACCGCATCGAGCTGAAAGGGCATTCGATTCGAAGAGGACGGCAATCTCTGGGGCAC	3109
QY	421	AAGCTGAGATGCAACTCAACAAGCCCAACAGTGTATATCATGGCCGCAACAGAGAAAGAC	480
Db	3110	AAGCTGAGATGCAACTCAACAAGCCCAACAGTGTATATCATGGCCGCAACAGAGAAAGAC	3169
QY	481	GGCATCAAGTGAACTTCGAAGATCCGCCACAACATGAGAGACGGCAAGCTGCAAGTCCGC	540
Db	3170	GGCATCAAGTGAACTTCGAAGATCCGCCACAACATGAGAGACGGCAAGCTGCAAGTCCGC	3229
QY	541	GACCATCAACGAGAGAAACACCCCATCGGGAGCGGGCCGTCGTCGGCCGCAACAACAC	600
Db	3230	GACCATCAACGAGAGAAACACCCCATCGGGAGCGGGCCGTCGTCGGCCGCAACAACAC	3289
QY	601	TACCTGAGCACCCAGATCCGCCCTTGAGCAAAAGACCCCAACAGAGACCGCATCAATGATC	660
Db	3290	TACCTGAGCACCCAGATCCGCCCTTGAGCAAAAGACCCCAACAGAGACCGCATCAATGATC	3349
QY	661	CTGCTGAGATTCCGTGACCGCCGCCGGGATCACTTCGGCATGACGAGCTGTACAAGAG	720
Db	3350	CTGCTGAGATTCCGTGACCGCCGCCGGGATCACTTCGGCATGACGAGCTGTACAAGAG	3409
QY	721	CTTAGCCATGGCTTCCGCCGCGAGGTGAGAGAGAGAGAT	759
Db	3410	AGCGGCTCGCACTTGAAGTGCACCTCGACGAATTCAT	3448

RESULT 9
AX801202
20001000
C410 hr
DWN
140000
DATE 13-OCT-2003

DEFINITION	Sequence 11 from Patent WO03056022.
ACCESSION	AX801202

```

VERSION      AX801202.1  GI:37654035
KEYWORDS
SOURCE       synthetic construct

```

ORGANISM
bacterial construct
artificial sequences.

Oxford Biomedica (UK) Limited (GB)

```
source
1..6418
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:33610"
```

```
ORIGIN
/ note="PONY vector sequence"
```

Query macula	85.0%	DB	128.0%	1.3e-82;
Best Local Similarity	96.8%		Pred. No.	1.3e-82;
Matches 735; Conservative	0;	Mismatches 24;	Indels 0;	Gaps 0;

0y 1 ATGTGACCAAGGGCGAGAGCTGTTTCAACGGGGTGTGTGCCATTCCTGTGTCACTGTGAC 60
 Db 2630 ATGTGTAGCAAGGGCGAGAGCTGTTTCAACGGGGTGTGTGCCATTCCTGTGTCACTGTGAC 2745

QY 61 GGCGACGTAAACGGCCACCAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCGCATGCCACCTAC 120

Dp 2750 GAGCAGCTTAAACGGCCCAAGTTTACGCTGTCCGGCGAGGGCGAGGGCGCATGCCACCTAC 2800

121 GCGAGCTGACCCCTGAAGTTTCATCTGCACCAACGCGAAGCTGCCGTGCCTGGCCACC 180

181 CTCGTGACCA^{CC}CTGACCTACGGCGTCA^{GT}GCTTCAGCGCGTACCCCGACCA^{CT}GAAG 240

241 CAGACGACTTCTCAAGTCGGCATGCCGAGGCTACGTCAGAGAGCGACCATCTTC 300

Db 2930 CAGACGACCTTCTTCAAGTCCGATCGGCAAGGCTACGTCCAGGAGCGACCATCTTC 2989
Qy 301 TTCAAGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGGGCAACCTTG 360
Db 2990 TTCAAGGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGGGCAACCTTG 3049
Qy 361 GTGAACGCGATCGAGCTGAAGGGGATGAGCTTCAAGGAGGAGCGGCAACATCTG99GAC 420
Db 3050 GTGAACGCGATCGAGCTGAAGGGGATGAGCTTCAAGGAGGAGCGGCAACATCTG99GAC 3109
Qy 421 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGACAAAGCAAGAAC 480
Db 3110 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGACAAAGCAAGAAC 3169
Qy 481 GGCATCAAGGTGAAGTTCAGATCCGCGCAACATCGAGAGCGGAGGCTGAGCTGCC 540
Db 3170 GGCATCAAGGTGAAGTTCAGATCCGCGCAACATCGAGAGCGGAGGCTGAGCTGCC 3229
Qy 541 GACCACTACAGCAAGAACCCCGATGGGCGGCGCGCGCGCTGCTGCGCGCAACAC 600
Db 3230 GACCACTACAGCAAGAACCCCGATGGGCGGCGCGCGCGCTGCTGCGCGCAACAC 3289
Qy 601 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGCGGATCACTGCTC 660
Db 3290 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGCGGATCACTGCTC 3349
Qy 661 CTGCTGAGTTCGAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAAG 720
Db 3350 CTGCTGAGTTCGAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAAG 3409
Qy 721 CTTAGCCATGAGCTCCGCGCGGAGGTGAGAGGAGAT 759
Db 3410 AGCGGCGCGACTCTAGAGTGCACCTGAGAGATTCAT 3448

RESULT 10
AX814838
LOCUS AX814838 6418 bp DNA linear PAT 05-DEC-2003
DEFINITION Sequence 12 from Patent W003064665.
ACCESSION AX814838
VERSION AX814838.1 GI:39104019
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 Radcliffe, P.O., Miskin, J.E., Wilkes, F.J. and Mitrophanous, K.A.
AUTHORS
TITLE Viral vector
JOURNAL Patent: WO 03064665-A 12 07-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source location/Qualifiers
1..6418
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Plasmid
Sequence"

ORIGIN
Query Match 84.8%; Score 720.6; DB 6; Length 6418;
Best Local Similarity 96.8%; Pred. No. 1.3e-82;
Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGTCCTCTGCTGAGCTGAGC 60
Db 2690 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGTCCTCTGCTGAGCTGAGC 2749
Qy 61 GGGCAGCTAAAGCGGCAAGATTGAGCGTGTCCGCGAGGGGAGAGCGAGTCCACTTAC 120
Db 2750 GGGCAGCTAAAGCGGCAAGATTGAGCGTGTCCGCGAGGGGAGAGCGAGTCCACTTAC 2809
Qy 121 GGGCAGCTAAAGCGGCAAGATTGAGCGAGTGTCCGCGAGGGGAGAGCGAGTCCACTTAC 180

Db 2810 GGCAGAGTGAACCTGAAAGTTCACTGTGACACACCGGAAAGTGTCCCGTGCCTTGGCCAC 2869
Qy 181 CTGTGAGCAACCTGAGCTTACGCGGTGAGTGTCTTACCGGCTAACCCGCAACATGAA 240
Db 2870 CTGTGAGCAACCTGAGCTTACGCGGTGAGTGTCTTACCGGCTAACCCGCAACATGAA 2929
Qy 241 CAGACGACCTTCTTCAAGTCCGCGATGCCCGAAGGCTACGTCAGAGGAGCGACCATCTTC 300
Db 2930 CAGACGACCTTCTTCAAGTCCGCGATGCCCGAAGGCTACGTCAGAGGAGCGACCATCTTC 2989
Qy 301 TTCAAGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGGGCAACCTTG 360
Db 2990 TTCAAGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGGGCAACCTTG 3049
Qy 361 GTGAACGCGATCGAGCTGAAGGGGATGAGCTTCAAGGAGGAGCGGCAACATCTG99GAC 420
Db 3050 GTGAACGCGATCGAGCTGAAGGGGATGAGCTTCAAGGAGGAGCGGCAACATCTG99GAC 3109
Qy 421 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGCAAGCAAGAAC 480
Db 3110 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGCAAGCAAGAAC 3169
Qy 481 GGCATCAAGGTGAAGTTCAGATCCGCGCAACATCGAGAGCGGAGGCTGAGCTGCC 540
Db 3170 GGCATCAAGGTGAAGTTCAGATCCGCGCAACATCGAGAGCGGAGGCTGAGCTGCC 3229
Qy 541 GACCACTACAGCAAGAACCCCGATGGGCGGCGCGCGCTGCTGCGCGCAACAC 600
Db 3230 GACCACTACAGCAAGAACCCCGATGGGCGGCGCGCGCTGCTGCGCGCAACAC 3289
Qy 601 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGCGGATCACTGCTC 660
Db 3290 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGCGGATCACTGCTC 3349
Qy 661 CTGCTGAGTTCGAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAAG 720
Db 3350 CTGCTGAGTTCGAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAAG 3409
Qy 721 CTTAGCCATGAGCTCCGCGCGGAGGTGAGAGGAGAT 759
Db 3410 AGCGGCGCGACTCTAGAGTGCACCTGAGAGATTCAT 3448

RESULT 11
AX590202/c
LOCUS AX590202 12789 bp DNA linear PAT 24-JAN-2003
DEFINITION Sequence 9 from Patent W002083888.
ACCESSION AX590202
VERSION AX590202.1 GI:27901286
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 Goossens, A. and Inz, D.
AUTHORS
TITLE The use of genes encoding membrane transporter pumps to stimulate
JOURNAL the production of secondary metabolites in biological cells
Patent: WO 02083888-A 9 24-OCT-2002;
Vlaams Internuiversiteit Instituut voor Biotechnologie vzw. (BE)
FEATURES
source location/Qualifiers
1..12789
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector pK7MG2D"

ORIGIN
Query Match 84.7%; Score 720; DB 6; Length 12789;
Best Local Similarity 100.0%; Pred. No. 1.4e-82;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGGGAGAGAGCTTACCGGGGTGTGTCCTCTGCTGAGCTGAGC 60

```

Db      5797  ATGTGACCAAGGGGAGAGAGCTGTTCAACCGGGGGTGGCCATCTGGTGAAGCTGAC 5738
Qy      61    GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGAGGGGATGCCACCTAC 120
Db      5737  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGAGGGGATGCCACCTAC 5678
Qy      121  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 180
Db      5677  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 5618
Qy      181  CTGTCGACCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 240
Db      5617  CTGTCGACCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 5558
Qy      241  CAGACGACCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 300
Db      5557  CAGACGACCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 5498
Qy      301  TTCAAGGACGAGGCACTAAGACCGGCGAGGGGATGCCACCTAC 360
Db      5497  TTCAAGGACGAGGCACTAAGACCGGCGAGGGGATGCCACCTAC 5438
Qy      361  GTGAACCGGCACTAAGACCGGCGAGGGGATGCCACCTAC 420
Db      5437  GTGAACCGGCACTAAGACCGGCGAGGGGATGCCACCTAC 5378
Qy      421  AAGCTGAGTCACTAAGACCGGCGAGGGGATGCCACCTAC 480
Db      5377  AAGCTGAGTCACTAAGACCGGCGAGGGGATGCCACCTAC 5318
Qy      481  GGCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 540
Db      5317  GGCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 5258
Qy      541  GACCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 600
Db      5257  GACCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 5198
Qy      601  TACCTGACGACCGGAGTCCGCGCTGAGCAAGACCGGAGGAGGAGGATGATGTC 660
Db      5197  TACCTGACGACCGGAGTCCGCGCTGAGCAAGACCGGAGGAGGAGGATGATGTC 5138
Qy      661  CTGCTGAGGTTGTCGACCGGCGGAGGATCACTCTGGCATGAGGAGGCTGTCAAGAG 720
Db      5137  CTGCTGAGGTTGTCGACCGGCGGAGGATCACTCTGGCATGAGGAGGCTGTCAAGAG 5078

RESULT 12
AX814844 7326 bp DNA linear PAT 05-DEC-2003
LOCUS      AX814844
DEFINITION Sequence 18 from Patent WO03064665.
ACCESSION  AX814844
VERSION     AX814844.1 GI:39104025
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
REFERENCE    1
AUTHORS      Radcliffe, P.O., Miskin, J.E., Wilkes, F.J. and Mitrophanous, K.A.
TITLE        Viral vector
JOURNAL      Patent: WO 03064665-A 18 07-AUG-2003;
              Oxford Biomedica (UK) Limited (GB)
              Location/Qualifiers
FEATURES
  source
    1..7326
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="Description of Artificial Sequence: Plasmid
    Sequence"

ORIGIN
Query Match      84.6%; Score 719.4; DB 6; Length 7326;

```

```

Best Local Similarity 97.9%; Pred. No. 1.8e-82;
Matches 729; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy      1    ATGTGACCAAGGGGAGAGCTGTTCAACCGGGGGTGGCCATCTGGTGAAGCTGAC 60
Db      3333  ATGTGACCAAGGGGAGAGCTGTTCAACCGGGGGTGGCCATCTGGTGAAGCTGAC 3392
Qy      61    GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 120
Db      3393  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 3452
Qy      121  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 180
Db      3453  GGCACGTAACCGGCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 3512
Qy      181  CTGTCGACCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 240
Db      3513  CTGTCGACCAAGCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 3572
Qy      241  CAGACGACCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 300
Db      3573  CAGACGACCTTCAAGCTGTCGGGCGAGGGGATGCCACCTAC 3632
Qy      301  TTCAAGGACGAGGCACTAAGACCGGCGAGGGGATGCCACCTAC 360
Db      3633  TTCAAGGACGAGGCACTAAGACCGGCGAGGGGATGCCACCTAC 3692
Qy      361  GTGAACCGGCACTAAGACCGGCGAGGGGATGCCACCTAC 420
Db      3693  GTGAACCGGCACTAAGACCGGCGAGGGGATGCCACCTAC 3752
Qy      421  AAGCTGAGTCACTAAGACCGGCGAGGGGATGCCACCTAC 480
Db      3753  AAGCTGAGTCACTAAGACCGGCGAGGGGATGCCACCTAC 3812
Qy      481  GGCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 540
Db      3813  GGCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 3872
Qy      541  GACCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 600
Db      3873  GACCATCAAGGTAAGCTTCAAGTCCGCAACATCGAGGAGGAGGCTGCGC 3932
Qy      601  TACCTGACGACCGGAGTCCGCGCTGAGCAAGACCGGAGGAGGAGGATGATGTC 660
Db      3933  TACCTGACGACCGGAGTCCGCGCTGAGCAAGACCGGAGGAGGAGGATGATGTC 3992
Qy      661  CTGCTGAGGTTGTCGACCGGCGGAGGATCACTCTGGCATGAGGAGGCTGTCAAGAG 720
Db      3993  CTGCTGAGGTTGTCGACCGGCGGAGGATCACTCTGGCATGAGGAGGCTGTCAAGAG 4052
Qy      721  CTGACCATGAGCTTCCGCGGAGGAG 745
Db      4053  AGCGCGCGGAGCTGACCTGACAG 4077

RESULT 13
AX814843 7327 bp DNA linear PAT 05-DEC-2003
LOCUS      AX814843
DEFINITION Sequence 17 from Patent WO03064665.
ACCESSION  AX814843
VERSION     AX814843.1 GI:39104024
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
REFERENCE    1
AUTHORS      Radcliffe, P.O., Miskin, J.E., Wilkes, F.J. and Mitrophanous, K.A.
TITLE        Viral vector
JOURNAL      Patent: WO 03064665-A 17 07-AUG-2003;
              Oxford Biomedica (UK) Limited (GB)
              Location/Qualifiers
FEATURES
  source
    1..7327

```

ORIGIN

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Plasmid
Sequence"

Query Match 84.6%; Score 719.4; DB 6; Length 7327;
Best Local Similarity 97.9%; Pred. No. 1.8e-82;
Matches 723; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 60
DB 3334 ATGGTGAAGCAAGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3393
QY 61 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 120
DB 3394 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3453
QY 121 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 180
DB 3454 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3513
QY 181 CTGTGTGACCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 240
DB 3514 CTGTGTGACCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3573
QY 241 CAGCAGCACTCTTCAAGTCCGCGCATGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 300
DB 3574 CAGCAGCACTCTTCAAGTCCGCGCATGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 3633
QY 301 TTCAAGGACGACGCGCACTCAAGAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 360
DB 3634 TTCAAGGACGACGCGCACTCAAGAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 3693
QY 361 GTTAAACCGCATGCACTGTAAGGCGCATGCTTCAAGGAGGAGCGCAACATCTGTGGGCGAC 420
DB 3694 GTTAAACCGCATGCACTGTAAGGCGCATGCTTCAAGGAGGAGCGCAACATCTGTGGGCGAC 3753
QY 421 AAGCTGAGTAACTCAACAGCAAGCAAGTCAATGAGCGGCGCAAGGAGGAGCGCAAGGAGGAG 480
DB 3754 AAGCTGAGTAACTCAACAGCAAGCAAGTCAATGAGCGGCGCAAGGAGGAGCGCAAGGAGGAG 3813
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGGAGCGGCAAGGAGGAGCGGCAAGGAGGAG 540
DB 3814 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGGAGCGGCAAGGAGGAGCGGCAAGGAGGAG 3873
QY 541 GACCACTAACGAGCAAGAACACCCCATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 3874 GACCACTAACGAGCAAGAACACCCCATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933
QY 601 TACCTGAGCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 660
DB 3934 TACCTGAGCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3993
QY 661 CTGTGTGAGTTCGTGACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 720
DB 3994 CTGTGTGAGTTCGTGACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 4053
QY 721 CTTAAGCATGGCTTCCCGCGGAGG 745
DB 4054 AGCGCGCGGAGCTTGAAGCTGAGG 4078

RESULT 14

AX268217
LOCUS AX268217 8528 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 11 from Patent WO0175135.
ACCESSION AX268217
VERSION AX268217.1 GI:16541474
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct

artificial sequences.

REFERENCE

1 Kingsman, A.J., Maden, M.B. and Corcoran, J.B.
Retinoic acid receptor beta-2, its antagonists, and gene therapy
vectors for the treatment of neurological disorders
Patent: WO 0175135-A 11 11-OCT-2001;
Oxford Biomedica (UK) Limited (GB)

JOURNAL

FEATURES

source
1..8528
Location/Qualifiers

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PONY86 5'cPPT POS deICTS EIAV vector genome
plasmid"

ORIGIN

Query Match 84.6%; Score 719.4; DB 6; Length 8528;
Best Local Similarity 97.2%; Pred. No. 1.8e-82;
Matches 732; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 60
DB 2868 ATGGTGAAGCAAGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 2927
QY 61 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 120
DB 2928 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 2987
QY 121 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 180
DB 2988 GGGGACGCTAAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3047
QY 181 CTGTGTGACCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 240
DB 3048 CTGTGTGACCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3107
QY 241 CAGCAGCACTCTTCAAGTCCGCGCATGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 300
DB 3108 CAGCAGCACTCTTCAAGTCCGCGCATGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 3167
QY 301 TTCAAGGACGACGCGCACTCAAGAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 360
DB 3168 TTCAAGGACGACGCGCACTCAAGAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTG 3227
QY 361 GTTAAACCGCATGCACTGTAAGGCGCATGCTTCAAGGAGGAGCGCAACATCTGTGGGCGAC 420
DB 3228 GTTAAACCGCATGCACTGTAAGGCGCATGCTTCAAGGAGGAGCGCAACATCTGTGGGCGAC 3287
QY 421 AAGCTGAGTAACTCAACAGCAAGCAAGTCAATGAGCGGCGCAAGGAGGAGCGCAAGGAGGAG 480
DB 3288 AAGCTGAGTAACTCAACAGCAAGCAAGTCAATGAGCGGCGCAAGGAGGAGCGCAAGGAGGAG 3347
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGGAGCGGCAAGGAGGAGCGGCAAGGAGGAG 540
DB 3348 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGGAGCGGCAAGGAGGAGCGGCAAGGAGGAG 3407
QY 541 GACCACTAACGAGCAAGAACACCCCATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 3408 GACCACTAACGAGCAAGAACACCCCATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3467
QY 601 TACCTGAGCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 660
DB 3468 TACCTGAGCAACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3527
QY 661 CTGTGTGAGTTCGTGACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 720
DB 3528 CTGTGTGAGTTCGTGACCGGCGGAGGCTGTTCAACCGGCGGTGTGCTCCATCTGTGTGAGCTGGAC 3587
QY 721 CTTAAGCATGGCTTCCCGCGGAGG 753
DB 3588 AGCGCGCGGAGCTTGAAGCTGAGG 3620

RESULT 15
AX801203
LOCUS AX801203 9952 bp DNA linear PAT 13-OCT-2003
DEFINITION Sequence 12 from Patent WO03056022.
ACCESSION AX801203
VERSION AX801203.1 GI:37654036
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Radcliffe, P., Mitrophanous, K. and Themis, M.
TITLE Transgenic organism
JOURNAL Patent: WO 03056022-A 12 10-JUL-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1. 9952
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PONY vector sequence"

ORIGIN

Query Match 84.6%; Score 719.4; DB 6; Length 9952;
Best Local Similarity 97.9%; Pred. No. 1.7e-82;
Matches 729; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
QY 1 ATGTGACGACAGGCGGAGAGCTGTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
DB 5959 ATGTGACGACAGGCGGAGAGCTGTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 6018
QY 61 GCGCAGCTAAACGCGCACAGTTCCGCTGTCGCGGAGGCGGAGGCGATGCCACTAC 120
DB 6019 GCGCAGCTAAACGCGCACAGTTCCGCTGTCGCGGAGGCGGAGGCGATGCCACTAC 6078
QY 121 GGCAGCTGACCTCTGAAGTTCTATCTGCAACACCGGCAAGCTGCCGTGCTGCGCCAC 180
DB 6079 GGCAGCTGACCTCTGAAGTTCTATCTGCAACACCGGCAAGCTGCCGTGCTGCGCCAC 6138
QY 181 CTGCTGACCACTCTGACCTTACGCGGTGCAAGTCTTCAAGCGCTACCCGACCATGTAAG 240
DB 6139 CTGCTGACCACTCTGACCTTACGCGGTGCAAGTCTTCAAGCGCTACCCGACCATGTAAG 6198
QY 241 CAGCAGACTTCTTCAAGTCCGCAATGCCGAAAGGCTACGTCAGAGAGCGACCATCTTC 300
DB 6199 CAGCAGACTTCTTCAAGTCCGCAATGCCGAAAGGCTACGTCAGAGAGCGACCATCTTC 6258
QY 301 TTCAAGGACGACGGAACCTAACAAGACCGCGCGGAGTGAAGTTGAGAGGCGACACCTTG 360
DB 6259 TTCAAGGACGACGGAACCTAACAAGACCGCGCGGAGTGAAGTTGAGAGGCGACACCTTG 6318
QY 361 GTGAACCGCATGAGCTGAAAGGGCATGCACTTCAAGGAGGACGCGCAATCTCTGGGGAC 420
DB 6319 GTGAACCGCATGAGCTGAAAGGGCATGCACTTCAAGGAGGACGCGCAATCTCTGGGGAC 6378
QY 421 AAGCTGAGTACACTTACAACGACCAACGCTTATATCATGCTGCGACAGCAAGAAAC 480
DB 6379 AAGCTGAGTACACTTACAACGACCAACGCTTATATCATGCTGCGACAGCAAGAAAC 6438
QY 481 GGCATCAAGGTGAACTTCAAGATCCGCAACAATCGAGGACGCGAGCGTGCAGCTCGCC 540
DB 6439 GGCATCAAGGTGAACTTCAAGATCCGCAACAATCGAGGACGCGAGCGTGCAGCTCGCC 6498
QY 541 GACCACTACCAAGCAACACCCCATCGGCGAGCGGCCGTGCTGCTGCGACCAACAC 600
DB 6499 GACCACTACCAAGCAACACCCCATCGGCGAGCGGCCGTGCTGCTGCGACCAACAC 6558
QY 601 TACCTGAGACCAAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGCATCAATGCTC 660
DB 6559 TACCTGAGACCAAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGCATCAATGCTC 6618
QY 661 CTGCTGAGATTGTAAGCGCGCGGGATCACTCTCGGATGAGCGAGCTGTACAAGAAAG 720
```

```
DB 6619 CTGCTGAGATTGTAAGCGCGCGGGATCACTCTCGGATGAGCGAGCTGTACAAGTAA 6678
QY 721 CTTAGCCATGCTTCCCGCGGAGG 745
DB 6679 AGCGCGCCGACCTCTAGCTTGCAGG 6703
```

Search completed: May 15, 2004, 04:33:26
Job time : 3704 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 19:55:16 ; Search time 436 Seconds
(without alignments)
8282.038 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atggtgagcaaggcagga.....tagcatcatgttagatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_29Jan04.*

1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001as.*
5: geneseq2001bs.*
6: geneseq2002s.*
7: geneseq2003as.*
8: geneseq2003bs.*
9: geneseq2003cs.*
10: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	850	100.0	850	AAZ32767	Aaz32767 Green flu
2	850	100.0	850	AAZ32767	Aac64428 EGFP-MODC
3	850	100.0	850	AAZ32767	Aa16105 EGFP-MODC
4	850	100.0	4278	ADRE3067	Aa35067 Plasmid p
5	850	100.0	4862	ACCA4691	Acc44691 Plasmid p
6	846	99.5	4163	AAA29134	Aaa29134 pNF-kappa
7	846	99.5	4269	AAA92002	Aaa92002 pCRE5-dEG
8	846	99.5	4300	AAA29135	Aaa29135 pCRE5-d2E
9	845	99.4	845	AB161141	AB161141 Enhanced
10	845	99.4	845	AB161141	AB161141 Enhanced
11	845	99.4	845	AB161141	AB161141 Enhanced
12	845	99.4	845	AB161141	AB161141 Enhanced
13	845	99.4	845	AB161141	AB161141 Enhanced
14	845	99.4	845	AB161141	AB161141 Enhanced
15	845	99.4	845	AB161141	AB161141 Enhanced
16	845	99.4	845	AB161141	AB161141 Enhanced
17	845	99.4	845	AB161141	AB161141 Enhanced
18	845	99.4	845	AB161141	AB161141 Enhanced
19	845	99.4	845	AB161141	AB161141 Enhanced
20	845	99.4	845	AB161141	AB161141 Enhanced
21	845	99.4	845	AB161141	AB161141 Enhanced
22	845	99.4	845	AB161141	AB161141 Enhanced
23	845	99.4	845	AB161141	AB161141 Enhanced

24	719	84.6	10292	4	AAE83084	AAE83084 EIAV expr
25	719	84.6	10384	4	AAE83082	AAE83082 EIAV expr
26	719	84.6	10384	4	AAE83083	AAE83083 EIAV expr
27	719	84.6	10385	5	AAH43714	AAH43714 PONY8.3G
28	718.6	84.5	761	3	AAE81941	AAE81941 Green flu
29	718.6	84.5	1190	6	AAE89070	AAE89070 DNA encod
30	718.6	84.5	1377	4	AAE81537	AAE81537 Ires-GFP
31	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
32	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
33	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
34	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
35	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
36	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
37	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
38	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
39	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
40	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
41	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
42	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
43	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
44	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP
45	718.6	84.5	1377	6	AAE81537	AAE81537 Ires-GFP

ALIGNMENTS

RESULT 1	AAZ32767	standard; cDNA; 850 BP.
ID	AAZ32767	
XX	AAZ32767	
AC	AAZ32767	
DT	15-SEP-2003	(revised)
DT	31-JUN-2000	(first entry)
XX		
DE	Green fluorescent protein/ornithine decarboxylase fusion gene.	
XX		
KW	Enhanced green fluorescent protein; EGFP; ornithine decarboxylase; MODC;	
KW	fusion; reporter; degradation domain; C-terminal; gene expression;	
KW	protein localization; fluorescence; short-lived; destabilized; half-life;	
KW	rapid degradation; PEST sequence; EGFP-MODC; EGFP-MODC422-461;	
KW	rapid turnover; determination; short-term event; repetitive event;	
KW	sensitive; developmental process; transient; periodic; cyclic;	
KW	expression; protein transport; circadian rhythm; kinetic; dynamic;	
KW	transcription; ds.	
XX		
OS	Aequorea victoria.	
OS	Mus sp.	
OS	Chimeric.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..846
FT		/*tag= a
FT		/product= "EGFP-MODC fusion protein"
XX		
XX	MO9954348-A1.	
XX	28-OCT-1999.	
XX		
XX	13-NOV-1998;	98MO-US024323.
XX	17-APR-1998;	98US-00062102.
XX		
XX	(CLON-) CLONTECH LAB INC.	
XX		
XX	L1 X, Kain S;	
XX		
XX	WPL; 2000-013225/01.	
XX	P-PSDB; AAY50142.	
XX		
PT	New fusion protein useful for studying cell localization, for studying	
PT	cell lineage and for assaying activation or deactivation of	


```

QY 1 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGTCCTTCGTGTGAGCTGGAC 60
DB 1 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGTCCTTCGTGTGAGCTGGAC 60
QY 61 GCGCAGCTAAACGCGCCACAAAGTTCAGCGTGTCCCGCGAGGGCGAGGGGAGTCCACTTAC 120
DB 61 GCGCAGCTAAACGCGCCACAAAGTTCAGCGTGTCCCGCGAGGGCGAGGGGAGTCCACTTAC 120
QY 121 GCGCAGCTGAGCCCTGAGTTCATCTGACCAACCGGCAAGCTGTCCCTGGGCGGACC 180
DB 121 GCGCAGCTGAGCCCTGAGTTCATCTGACCAACCGGCAAGCTGTCCCTGGGCGGACC 180
QY 181 CTGCTGACCACTTGAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 240
DB 181 CTGCTGACCACTTGAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 240
QY 241 CAGCAGCACTTCTTCAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 300
DB 241 CAGCAGCACTTCTTCAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 300
QY 301 TTCAAGAGCAAGCGGCACTACAAAGCCGCGCGGAGTGAAGTTCAGAGGCGGACCCCTG 360
DB 301 TTCAAGAGCAAGCGGCACTACAAAGCCGCGCGGAGTGAAGTTCAGAGGCGGACCCCTG 360
QY 361 GTGAACCGCATTCAGCTGAGAGGCGCATCTTCAAGAGGAGCGGACATCTCGGCGAC 420
DB 361 GTGAACCGCATTCAGCTGAGAGGCGCATCTTCAAGAGGAGCGGACATCTCGGCGAC 420
QY 421 AAGCTGAGTCACTCACTACAAAGCCCAACGCTTATCATGAGCGGCAAGAGAGAAC 480
DB 421 AAGCTGAGTCACTCACTACAAAGCCCAACGCTTATCATGAGCGGCAAGAGAGAAC 480
QY 481 GGCATCAAGTGAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 540
DB 481 GGCATCAAGTGAAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 540
QY 541 GACCACTACAGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 600
DB 541 GACCACTACAGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 600
QY 601 TACCTGAGCAAGCGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 660
DB 601 TACCTGAGCAAGCGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 660
QY 661 CTGCTGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 720
DB 661 CTGCTGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 720
QY 721 CTGCTGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 780
DB 721 CTGCTGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 780
QY 781 TGTGCGCAGAGAGCGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 840
DB 781 TGTGCGCAGAGAGCGGAGTTCAGTGTGAGTGTCTTCAAGCCGCTACCCGACCAATGAG 840
QY 841 GTGTAGATGC 850
DB 841 GTGTAGATGC 850

```

RESULT 4

ADE35067 standard; DNA; 4278 BP.

ADE35067;

29-JAN-2004 (first entry)

Plasmid pΔ2EGFP-1.

stem cell; green fluorescent protein; pΔ2EGFP; gene; ds.

```

OS Chimeric.
OS Aequorea victoria.
PN W02003080816-A2.
PD 02-OCT-2003.
PF 18-MAR-2003; 2003MO-GB001111.
PR 19-MAR-2002; 2002GB-00006422.
PR 08-MAY-2002; 2002GB-00010458.
XX (UYSH-) UNIV SHEFFIELD.
XX Andrews P, Draper J;
XX WPI, 2003-779256/73.
DR Manipulating phenotype of stem cell by providing cell transfected with
XX nucleic acid comprising promoter which confers substantial stem cell
XX specific expression on selective marker gene(s), and proliferating cell.
XX Disclosure; Fig 5; 40pp; English.
XX
XX The present sequence is that of plasmid pΔ2EGFP-1 comprising an enhanced
XX green fluorescent protein (GFP) reporter gene. This expression vector was
XX used to demonstrate a method of manipulating the phenotype of a stem
XX cell. In this method, a cell is transfected with a nucleic acid molecule
XX that includes a promoter which comprises a motif ADE35058-ADE35060 that
XX confers stem cell specific expression on a selectable marker (e.g.
XX fluorescent protein or protein activating polypeptide) gene, and
XX conditions conducive to proliferation of the cell are provided. A cell
XX culture system is provided which facilitates the maintenance of stem
XX cells, particularly embryonic stem cells, in an undifferentiated state.
XX Also provided are differentiated cells and tissues. The genome of which
XX includes a nucleic acid construct comprising a promoter which has a stem
XX cell specific expression pattern which controls expression of a gene the
XX expression of which allows the selective ablation of cells which have de-
XX differentiated to a stem cell phenotype, thereby allowing their removal
XX from a population of differentiated cells. The stem cells are
XX haematopoietic, neural, bone, muscle, mesenchymal, trophoblastic,
XX epithelial, endodermal or embryonic stem cells (including embryonal
XX carcinoma TERA2 or NTera2 cells), or embryonal germ cells.
XX
XX Sequence 4278 BP; 1015 A; 1177 C; 1150 G; 936 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 850; DB 9; Length 4278;
XX Best Local Similarity 100.0%; Pred. No. 1,2e-138;
XX Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 361 GTGAACCGCATGAGCTGAAAGGCGATCTCAAGAGGAGGCAACATCTGGGGCAC 420
XX
DB 457 GTGAACCGCATGAGCTGAAAGGCGATCTCAAGAGGAGGCAACATCTGGGGCAC 516
QY 421 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGGCCGCAAGCAAGAAAC 480
DB 517 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGGCCGCAAGCAAGAAAC 576
QY 481 GGATCAAGTGAATCTCAAGATCCGCCCAACATCGAGGAGCGGAGCTGAGCTGCC 540
DB 577 GGATCAAGTGAATCTCAAGATCCGCCCAACATCGAGGAGCGGAGCTGAGCTGCC 636
QY 541 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGAGCTGAGCTGCCGCAACAC 600
DB 637 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGAGCTGAGCTGCCGCAACAC 696
QY 601 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGAGGCGCATCATGTC 660
DB 697 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGAGGCGCATCATGTC 756
QY 661 CTGCTGAGTTCGTCAGCGCGCGCGGATCACTCTGCGCATGAGCAAGCTGTACAAAG 720
DB 757 CTGCTGAGTTCGTCAGCGCGCGCGGATCACTCTGCGCATGAGCAAGCTGTACAAAG 816
QY 721 CTGAGCATGCTTCCCGCGAGGTGAGAGGAGATGAGCAAGCTGCCATGTCT 780
DB 817 CTGAGCATGCTTCCCGCGAGGTGAGAGGAGATGAGCAAGCTGCCATGTCT 876
QY 781 TGTGCGCAGAGGAGCGGAGTGAACCCCTGACCCCTGCTGCTTGTGTAAGATCAAT 840
DB 877 TGTGCGCAGAGGAGCGGAGTGAACCCCTGACCCCTGCTGCTTGTGTAAGATCAAT 936
QY 841 GTGTAGATGC 850
DB 937 GTGTAGATGC 946

RESULT 5
ACCA4691
ID ACCA4691 standard; DNA; 4862 BP.
XX
ACCA4691;

29-MAY-2003 (first entry)

Plasmid pD2eGFP-N1 nucleotide fragment SEQ ID NO:87.

XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;
KM att site; integrase; recombinase; Aces; gene therapy; transgenic animal;
XX Platform artificial chromosome expression system; gene; ds.
OS Synthetic.

PN WO200297059-A2.

PD 05-DEC-2002.

PF 30-MAY-2002; 2002WO-US017452.

PR 30-MAY-2001; 2001US-0294758P.

PR 21-MAR-2002; 2002US-0366891P.

XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

PA Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;
PI Stewart S, Sheldahl J;

DR WPI; 2003-140461/13.

XX Novel eukaryotic chromosome comprising one or many att sites which
PT permits site-directed integration in the presence of lambda-integrase,
PT useful for site-specific recombination-directed integration of DNA of

PT interest.

PS Example 2; Page 209-210; 272pp; English.

XX The present invention describes a eukaryotic chromosome (I) comprising
CC one or several att sites, where an att site is heterologous to the
CC chromosome, and permits site-directed integration in the presence of
CC lambda-integrase. Also described: (1) a platform artificial chromosome
CC expression system (Aces) (II) comprising several att sites that participate
CC in recombinase catalyzed recombination; and (2) a method (M1) for
CC introducing a heterologous nucleic acid into a platform artificial
CC chromosome. (I) can be used in gene therapy. (M1) is useful for
CC introducing a heterologous nucleic acid molecule into a platform
CC artificial chromosome, preferably an Aces. (II) is useful for producing a
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection
CC by a carrier system, microinjection, microcell fusion, electroporation,
CC microprojectile bombardment or direct DNA transfer into a heterologous
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous
CC nucleic acid that encodes a therapeutic product which is useful for
CC making a library of Aces comprising random portions of a genome. ACCA4612
CC to ACCA4732 and ABP96650 to ABP96657 represent sequences used in the
CC exemplification of the present invention

XX Sequence 4862 BP; 1162 A; 1291 G; 1089 T; 0 U; 0 Other;

Query Match 100.0%; Score 850; DB 7; Length 4862;

Best Local Similarity 100.0%; Pred. No. 1.2e-138; Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTGAAGCAAGGCGAGAGCTGTTACCGGGGTTGTTCCATCTGTGAGCTGAGAC 60
DB 679 AAGTGAAGCAAGGCGAGAGCTGTTACCGGGGTTGTTCCATCTGTGAGCTGAGAC 738
QY 61 GGCGAGCTAAACGGCCCAAGTTACAGCTGTCCGGGAGAGGCGAGCGATGCACTTAC 120
DB 739 GGCGAGCTAAACGGCCCAAGTTACAGCTGTCCGGGAGAGGCGAGCGATGCACTTAC 798
QY 121 GGCAAGCTACCTGTAAGTTCTGACACCAACCGGCAAGCTGCCGCTGAGCCCAAC 180
DB 799 GGCAAGCTACCTGTAAGTTCTGACACCAACCGGCAAGCTGCCGCTGAGCCCAAC 858
QY 181 CTCTGTAACCACTCTGACCTTACAGGCTGTGAGCTGTTACCTCCAGCAATGAAAG 240
DB 859 CTCTGTAACCACTCTGACCTTACAGGCTGTGAGCTGTTACCTCCAGCAATGAAAG 918
QY 241 CAGCAGACTTCTTCAAGTCCGCGATGCCGCAAGGCTAGTCCAGAGGCGCAATCTTC 300
DB 919 CAGCAGACTTCTTCAAGTCCGCGATGCCGCAAGGCTAGTCCAGAGGCGCAATCTTC 978
QY 301 TTCAAGAGCAGCGCAACTTCAAGACCCGCGGAGGTGAAGTTCCAGGCGCAACCTTG 360
DB 979 TTCAAGAGCAGCGCAACTTCAAGACCCGCGGAGGTGAAGTTCCAGGCGCAACCTTG 1038
QY 361 GTGAACCGCATGAGCTGAAAGGCGATCTCAAGAGGAGGCAACATCTGGGGCAC 420
DB 1039 GTGAACCGCATGAGCTGAAAGGCGATCTCAAGAGGAGGCAACATCTGGGGCAC 1098
QY 421 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGGCCGCAAGCAAGAAAC 480
DB 1099 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGGCCGCAAGCAAGAAAC 1158
QY 481 GGATCAAGTGAATCTCAAGATCCGCCCAACATCGAGGAGCGGAGCTGAGCTGCC 540
DB 1159 GGATCAAGTGAATCTCAAGATCCGCCCAACATCGAGGAGCGGAGCTGAGCTGCC 1218
QY 541 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGAGCTGAGCTGCCGCAACAC 600
DB 1219 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGAGCTGAGCTGCCGCAACAC 1278
QY 601 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGAGGCGCATCATGTC 660
DB 1279 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGAGGCGCATCATGTC 1338

KW CAMP level; signal transduction; CAMP response element; CRE;
 KW destabilised enhanced green fluorescent protein; d2EGFP;
 KW reporter construct; ds.

OS Synthetic.

EN WO20049161-A1.

XX 24-AUG-2000.

XX 17-FEB-2000; 2000WO-US004165.

XX 17-FEB-1999; 99US-0120463P.

XX (CLON-) CLONTECH LAB INC.

XX Li X, Zhao X;

XX WPI; 2000-558302/51.

XX Novel reporter cassette for monitoring cyclic adenosine monophosphate
 PT levels, comprises cyclic adenosine monophosphate response elements and a
 PT gene encoding a fluorescent reporter molecule.

XX Example 4; Page 16-18; 29pp; English.

XX The present sequence is the sequence of a reporter construct used in the
 CC methods of the invention to monitor cyclic adenosine monophosphate (CAMP)
 CC levels. It comprises 5 copies of the CAMP response elements (CREs), the
 CC gonadotrophin alpha-gene promoter and the destabilised enhanced green
 CC fluorescent protein (d2EGFP) gene. The reporter construct can be used in
 CC high-throughput assays to screen for factors and mutants involved in the
 CC CAMP signal transduction pathway or to monitor CAMP levels in real time
 CC in vivo

XX Sequence 4269 BP; 1050 A; 1135 C; 1057 G; 1027 T; 0 U; 0 Other;

XX Query Match 99.5%; Score 846; DB 3; Length 4269;

XX Best Local Similarity 100.0%; Pred. No. 5.9e-138;

XX Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGACCAAGGGGAGAGAGTGTTCACCGGGGTGGTCCATCTGTGAGCTGGAC 60
 DB 347 ATGGTGACCAAGGGGAGAGAGTGTTCACCGGGGTGGTCCATCTGTGAGCTGGAC 406
 QY 61 GGGCAGCTAAACGGCCACAGATTGACGCTGTCGGGAGGGGAGGGGAGATGCCACTAC 120
 DB 407 GGGCAGCTAAACGGCCACAGATTGACGCTGTCGGGAGGGGAGGGGAGATGCCACTAC 466
 QY 121 GGGCAGCTGACCTGTAAGTTCACTGACCAACCGGCAAGCTGCGCTGCGCCCAAC 180
 DB 467 GGGCAGCTGACCTGTAAGTTCACTGACCAACCGGCAAGCTGCGCTGCGCCCAAC 526
 QY 181 CTCGTGACCAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 240
 DB 527 CTCGTGACCAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 586
 QY 241 CAGCAGCACTTCTTCAAGTCCGATGCGCCGAGGCTACGTCACGAGACGACACATCTTC 300
 DB 587 CAGCAGCACTTCTTCAAGTCCGATGCGCCGAGGCTACGTCACGAGACGACACATCTTC 646
 QY 301 TTCAGAGCAAGCGGCACTACCAAGACCGGCGCGAGGTGAAGTTGAGGGGACACCTTG 360
 DB 647 TTCAGAGCAAGCGGCACTACCAAGACCGGCGCGAGGTGAAGTTGAGGGGACACCTTG 706
 QY 361 GTGAACCGATGAGCTGAGAGGAGCATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 707 GTGAACCGATGAGCTGAGAGGAGCATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 766
 QY 421 AAGGTGAGTACCACTACCAAGACCAACGCTCTATATCATGAGCCGACAGAGAGAGAG 480
 DB 767 AAGGTGAGTACCACTACCAAGACCAACGCTCTATATCATGAGCCGACAGAGAGAGAG 826

QY 481 GGCATCAAGTGAACTTCAAGATCCGCCACACATGAGAGAGAGAGAGAGAGAGAGAGAG 540
 DB 827 GGCATCAAGTGAACTTCAAGATCCGCCACACATGAGAGAGAGAGAGAGAGAGAGAGAG 886
 QY 541 GACCACTACCAAGAGAACACCCCATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 DB 887 GACCACTACCAAGAGAACACCCCATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
 QY 601 TACCTGAGACCCAGTCCGCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 947 TACCTGAGACCCAGTCCGCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
 QY 661 CTGCTGAGATTGTAACCGCGCGCGGAGATCACTCTGCGCATGAGAGAGAGAGAGAGAG 720
 DB 1007 CTGCTGAGATTGTAACCGCGCGCGGAGATCACTCTGCGCATGAGAGAGAGAGAGAGAG 1066
 QY 721 CTGAGCAATGCTTCCCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 DB 1067 CTGAGCAATGCTTCCCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
 QY 781 TGTGCCAG 840
 DB 1127 TGTGCCAG 1186
 QY 841 GTGTAG 846
 DB 1187 GTGTAG 1192

RESULT 8
 AAA29135
 ID AAA29135 standard; DNA; 4300 BP.
 AC AAA29135;
 XX 12-SEP-2000 (first entry)
 DT PCR5-d2EGFP construct.
 DE Vector; transcription; cis-element; CAMP responsive element 5; CRE5;
 KW d2EGFP; red-shifted; destabilized; green fluorescent protein; reporter;
 KW promoter; monitor; signal transduction; ss.
 OS Synthetic.
 XX WO200034435-A2.
 XX 15-JUN-2000.
 PD 01-DEC-1999; 99WO-US028451.
 PF 08-DEC-1998; 98US-00206887.
 PR (CLON-) CLONTECH LAB INC.
 PA Li X, Zhao X, Jiang X, Fang Y, Duong T, Kain S;
 XX WPI; 2000-423400/36.
 DR Cis element-reporter construct for measuring transcription, comprises a
 PT reporter gene, promoter and a cis element such as nuclear factor-kappaB,
 PT heat responsive element, Myc, p53 or alkaline phosphatase 1.
 PS Example 6; Page 37-39; 42pp; English.
 XX A new construct for measuring transcription comprises a cis-element from
 CC a DNA sequence such as that for nuclear factor-kappaB (NF-kappaB), heat
 CC responsive element (HRE), Myc, p53, nuclear factor of activated T-cells
 CC (NFAT), activator protein 1 (AP-1), serum responsive element (SRE),
 CC glucocorticoid responsive element (GRE) or cyclic adenosine monophosphate
 CC (CAMP) responsive element, and a reporter gene and a promoter. Reporter
 CC genes are chosen from secreted alkaline phosphatase (SEAP), destabilized
 CC green fluorescent protein (d2EGFP) and luciferase (Luc). Promoter are

CC preferably chosen from the thymidine kinase and gonadotropin-alpha gene
 CC promoters. The constructs are useful for measuring transcription and for
 CC monitoring the activation of a transcription factor. They are also used
 CC to determine whether a gene is involved in signal transduction pathway,
 CC where the cis-element is a component of a DNA sequence which affects the
 CC expression of a test gene

SQ Sequence 4300 BP; 1058 A; 1139 C; 1065 G; 1038 T; 0 U; 0 Other;

Query Match 99.5%; Score 846; DB 3; Length 4300;

Best Local Similarity 100.0%; Pred. No. 5.9e-138; Mismatches 0; Gaps 0;

Matches 846; Conservative 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCAAGAGGCGAGAGCTGTTCAACCGGGGTGTGCCCATCTGCTGAGCTGAC 60
DB 378 ATGTGAGCAAGAGGCGAGAGCTGTTCAACCGGGGTGTGCCCATCTGCTGAGCTGAC 437
QY 61 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 120
DB 438 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 497
QY 121 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 180
DB 498 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 557
QY 181 CTGTGACCAACCTGACCTACCGGCGTGCAGTGTCTTACCGCTTACCGGCGATCAAG 240
DB 558 CTGTGACCAACCTGACCTACCGGCGTGCAGTGTCTTACCGCTTACCGGCGATCAAG 617
QY 241 CAGCAGATTTCTTCAAGTCCGCGATGCCGGAAGGTACAGTCCAGAGAGGCGACATCTTC 300
DB 618 CAGCAGATTTCTTCAAGTCCGCGATGCCGGAAGGTACAGTCCAGAGAGGCGACATCTTC 677
QY 301 TTCAAGAGCAGCGGCGCACTACAAGACCGCGCGCGAGGTGAAGTTGAGGGCGACACCTG 360
DB 678 TTCAAGAGCAGCGGCGCACTACAAGACCGCGCGCGAGGTGAAGTTGAGGGCGACACCTG 737
QY 361 GTGAACCGCATGAGCTGAGGAGGCGATCATCTTCAAGAGAGCAGGCAACATCTCTGGGCGAC 420
DB 738 GTGAACCGCATGAGCTGAGGAGGCGATCATCTTCAAGAGAGCAGGCAACATCTCTGGGCGAC 797
QY 421 AAGCTGAGATTAACATCAAGCAAGCCCAACCGTCTATATCATGCGCGCAAGAGAAAGAC 480
DB 798 AAGCTGAGATTAACATCAAGCAAGCCCAACCGTCTATATCATGCGCGCAAGAGAAAGAC 857
QY 481 GGCATCAAGGTAACTTCAAGATCCGCGCAACATCGAGAGAGCGGCGAGCTGCGCC 540
DB 858 GGCATCAAGGTAACTTCAAGATCCGCGCAACATCGAGAGAGCGGCGAGCTGCGCC 917
QY 541 GACCACTACGAGAGAACCCCCCATGCGGAGCGGCGCGTGTCTGCTGCCGCAACACAC 600
DB 918 GACCACTACGAGAGAACCCCCCATGCGGAGCGGCGCGTGTCTGCTGCCGCAACACAC 977
QY 601 TACCTGAGACCAACGTCGCGCTGAGCAAAAGACCCCAAGAGAGGCGATCAAGTGTG 660
DB 978 TACCTGAGACCAACGTCGCGCTGAGCAAAAGACCCCAAGAGAGGCGATCAAGTGTG 1037
QY 661 CTGTGAGATTTGATGACCGCGCGCGGAGTCACTCTGCGATGAGCAGAGCTGTACAAGAG 720
DB 1038 CTGTGAGATTTGATGACCGCGCGCGGAGTCACTCTGCGATGAGCAGAGCTGTACAAGAG 1097
QY 721 CTTAAGCATAGGCTCTCCCGCGGAGGTGAGAGAGAGAGATGATGACCGCTGCCATGTCT 780
DB 1098 CTTAAGCATAGGCTCTCCCGCGGAGGTGAGAGAGAGAGATGATGACCGCTGCCATGTCT 1157
QY 781 TGTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 1158 TGTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
QY 841 GTGTAG 846
DB 1218 GTGTAG 1223

```

RESULT 9

AB161141
 ID ABR61141 standard; DNA; 845 BP.

XX ABR61141;

XX 18-SEP-2002 (first entry)

XX Enhanced green fluorescent protein reporter DNA fragment.

XX Alpha-SMA, smooth muscle actin; screening; smooth muscle cell; EGFP;
 XX myofibroblast; gene therapy; enhanced green fluorescent protein; ds.

XX Synthetic.

XX EPI172375-A1.

XX 16-JAN-2002.

XX 22-DEC-2000; 2000EP-00128446.

XX 11-JUN-2000; 2000DE-01033633.

XX 31-OCT-2000; 2000DE-01053879.

XX (ODEN/) ODENTHAL M.

XX Odenthal M, Jung D;

XX WPI; 2002-149590/20.

XX New nucleic acid containing regulatory region of the smooth muscle actin

XX gene, useful e.g. for manipulating gene expression in smooth muscle

XX cells.

XX PS

XX XX

XX This invention describes a novel nucleic acid (I) comprising: (i) at

XX least one functional region (Ia) from the regulatory region of the alpha-

XX smooth muscle actin (SMA) gene and (ii) at least one additional

XX functional sequence (Ib) operably linked to (Ia). The products of the

XX invention can be used for preparing genetically modified eukaryotic cells

XX or organisms, for isolation and screening of smooth muscle cells,

XX myofibroblasts or related cells, and for manipulation of gene expression

XX and/or cell function in smooth muscle cell or myofibroblasts,

XX particularly for gene therapy. Component (Ia) provides cell-type or

XX differentiation-specific expression or modulation of genes. This sequence

XX represents a DNA fragment of destabilised enhanced green fluorescent

XX protein (EGFP) which can be used as a reporter molecule under the control

XX of the alpha-SMA described in the invention

XX CC

XX SQ

XX Sequence 845 BP; 196 A; 274 C; 246 G; 129 T; 0 U; 0 Other;

Query Match 99.4%; Score 845; DB 6; Length 845;

Best Local Similarity 100.0%; Pred. No. 8.8e-138; Mismatches 0; Gaps 0;

Matches 845; Conservative 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCAAGAGGCGAGAGCTGTTCAACCGGGGTGTGCCCATCTGCTGAGCTGAC 60
DB 1 ATGTGAGCAAGAGGCGAGAGCTGTTCAACCGGGGTGTGCCCATCTGCTGAGCTGAC 60
QY 61 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 120
DB 61 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 120
QY 121 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 180
DB 121 GGCGACGTAAACGGCCCAAGTTCAAGCGTGTCCGCGAGGGCGAGGGCGATGCCCTAC 180
QY 181 CTGTGACCAACCTGACCTACCGGCGTGCAGTGTCTTACCGCTTACCGGCGATCAAG 240
DB 181 CTGTGACCAACCTGACCTACCGGCGTGCAGTGTCTTACCGCTTACCGGCGATCAAG 240

```


XX	RESULT 12
XX	ACF04133
ID	ACF04133 standard; DNA; 6418 BP.
XX	ACF04133;
XX	15-OCT-2003 (first entry)
DT	
XX	Equine infectious anaemia virus vector pONY8-1G.
DE	
XX	Transgenic; cell; lentiviral expression vector; viral vector; aptazyme;
KW	vector; ds.
XX	
OS	Equine infectious anaemia virus.
XX	
PN	WO2003056022-A2.
XX	
PD	10-JUL-2003.
XX	
PF	23-DEC-2002; 2002WO-GB005901.
XX	
PR	21-DEC-2001; 2001GB-00030797.
PR	18-JAN-2002; 2002GB-00001140.
PR	26-FEB-2002; 2002US-00082122.
PR	17-MAY-2002; 2002GB-00011409.
XX	
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.
XX	
PI	Radeliffe P, Mitrophanous K, Themis M;
XX	
DR	WPI; 2003-569455/53.
XX	
PT	Producing a transgenic cell comprises introducing into a cell a non-primate lentiviral expression vector comprising a nucleotide of interest.
XX	
PS	Example 2; Fig 13; 117pp; English.
CC	The present invention relates to a method of producing a transgenic cell comprising introducing into a cell a non-primate lentiviral expression vector comprising a nucleotide of interest (NOI). The method is useful for producing a transgenic cell. The present sequence is a vector sequence used in the exemplification of the invention
CC	
SQ	Sequence 6418 BP; 1673 A; 1559 C; 1570 G; 1616 T; 0 U; 0 Other;
	Query Match 84.8%; Score 720.6; DB 8; Length 6418;
	Best Local Similarity 96.8%; Pred. No. 3.3e-116f;
	Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY	1 ATGTGAGCAGAAGCGAGAGCTTTCACCGGGGTGTGCCATCTTGAGAGTTGAC 60
DB	2690 ATGTGAGCAGAAAGCGAGAGCTGTTCACCGGGGTGTGCCATCTTGAGAGTTGAC 2749
QY	61 GGCGACGTAAAGGGCCACAAGTTCAAGCGTGTCCGGCGAAGGGCGAAGGCACCTAC 120
DB	2750 GGCGACGTAAAGGGCCACAAGTTCAAGCGTGTCCGGCGAAGGGCGAAGGCACCTAC 2809
QY	121 GGCAAGCTGACCCTGAAGTTCAATGTGCACCAACGGGCAAGCTGCCGTGCTGSCCAAC 180
DB	2810 GGCAAGCTGACCCTGAAGTTCAATGTGCACCAACGGGCAAGCTGCCGTGCTGSCCAAC 2869
QY	181 CTCGTGACCAACCTTGACCTTAAGCGGCGTGCAGTGTCTTAGCGGTTACCCGACCAATGAG 240
DB	2870 CTCGTGACCAACCTTGACCTTAAGCGGCGTGCAGTGTCTTAGCGGTTACCCGACCAATGAG 2922
QY	241 CAGCAGCACTTTCTTAAGTCGCGCATGCCGGAAGGCTACGTCCAGAGCGCACACTTTC 300
DB	2930 CAGCAGCACTTTCTTAAGTCGCGCATGCCGGAAGGCTACGTCCAGAGCGCACACTTTC 2988
QY	301 TTCAGGACGACGGCACTTCAAGAACCCGGCGCGAGGTGAATTGAGGGCGACACCTTG 360

CC This invention relates to a novel multicistronic retroviral vector genome
CC comprising a first nucleic acid sequence upstream of an integral
CC regulatory element, such that the level of genomic RNA available for
CC packaging in the absence of rev, or its functional equivalent, is
CC increased. Specifically, the vector is a lentiviral genome that comprises
CC a therapeutically important nucleotide sequence of interest operably
CC linked to a tetracycline operator. Furthermore, the first nucleic acid
CC sequence represents a selectable marker such as the tetracycline
CC repressor (Tet^r) gene, which is codon optimised for expression in
CC mammalian cells. Together, the retroviral vector particle, DNA construct
CC and mammalian cell are useful for the preparation of a medicament to
CC deliver the therapeutic sequence of interest to a target site. As such,
CC the present invention describes a method that can be used in the
CC treatment of inflammatory disease, asthma, psoriasis, cancer e.g. chronic
CC myeloid leukaemia (CML) and neurodegenerative disease e.g. Alzheimer's.
CC Accordingly, via gene therapy these vectors can be described as having
CC viruloid, neurotropic, neuroprotective, immunosuppressive and/or
CC anti-inflammatory activity. This polynucleotide sequence represents the
CC total plasmid DNA of pONY8.1G of the invention.

CC Sequence 6418 BP; 1673 A; 1559 C; 1570 G; 1616 T; 0 U; 0 Other;

Query Match 84.8%; Score 720.6; DB 8; Length 6418;

Best Local Similarity 96.8%; Pred. No. 3.3e-116; Indels 0; Gaps 0;

Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGGTAGAGCAAGGGGAGGAGGCTGTTCAACGGGGTGGTGGCCATCTGCTGAGCTGAGAC 60
DB 2690 ATGGTAGAGCAAGGGGAGGAGGCTGTTCAACGGGGTGGTGGCCATCTGCTGAGCTGAGAC 2749
QY 61 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGCGAGGCGAGGCGAGTGCACCTAC 120
DB 2750 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGCGAGGCGAGGCGAGTGCACCTAC 2809
QY 121 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGCGAGGCGAGGCGAGTGCACCTAC 180
DB 2810 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGCGAGGCGAGGCGAGTGCACCTAC 2869
QY 181 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 240
DB 2870 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2929
QY 241 CAGCAGCACTTCTTCAAGTCCGCGATGCCGAGGCTTACCTGCAAGGCGACCACTTTC 300
DB 2930 CAGCAGCACTTCTTCAAGTCCGCGATGCCGAGGCTTACCTGCAAGGCGACCACTTTC 2989
QY 301 TTGAGAGCAAGCGGCACTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 360
DB 2990 TTGAGAGCAAGCGGCACTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 3049
QY 361 GTGAGACCGATCGAGCTGAGGCGATCGATCGATCGATCGATCGATCGATCGATCGAT 420
DB 3050 GTGAGACCGATCGAGCTGAGGCGATCGATCGATCGATCGATCGATCGATCGATCGAT 3109
QY 421 AAGCTGAGTGAACCTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 480
DB 3110 AAGCTGAGTGAACCTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 3169
QY 481 GGGATCAAGGTAAGCTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 540
DB 3170 GGGATCAAGGTAAGCTTCAAGGCGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAG 3229
QY 541 GACCACTACGAGCAACACCCCATCGGCGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 600
DB 3230 GACCACTACGAGCAACACCCCATCGGCGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 3289
QY 601 TACCTGAGCAACCGAGTCCGCTTGAAGCAAGACCCCAAGAGAGGCGGCGAGTCAATGTC 660
DB 3290 TACCTGAGCAACCGAGTCCGCTTGAAGCAAGACCCCAAGAGAGGCGGCGAGTCAATGTC 3349
QY 661 CTGCTGAGTTCGAGCGCGCGCGGAGTCACTCTCGGCAATGAGCAAGCTGTACAAAG 720
DB 3350 CTGCTGAGTTCGAGCGCGCGCGGAGTCACTCTCGGCAATGAGCAAGCTGTACAAAG 3409

QY 721 CTAGCCATGAGCTTCCGCGAGGAGTGAAGAGCAAGAT 759
DB 3410 AGCGCGCGGAGCTCTAAGAGTGAACCTGAGGAAATTCAT 3448

RESULT 14
ID ABS57536/C
ABS57536 standard; cDNA; 12789 BP.

AC ABS57536;

DT 10-FEB-2003 (first entry)

DE Vector pK7MG2D DNA.

KW PDR1; pleiotropic drug resistance; ATP-binding cassette-transporter;

KW ABC-transporter; secondary metabolite production; transgenic plant;

KW plant disease resistance; secondary metabolite secretion; plant;

KW pest resistance; anti-oxidative; anti-metastatic; alkaloid; cytostatic;

KW terpenoid indole alkaloid tropone alkaloid; steroid alkaloid;

KW polyhydroxy alkaloid; pesticide; vector; ds.

OS Synthetic.

PN WO20028388-A2.

PP 18-OCT-2002.

PR 18-APR-2002; 2002MO-EP004322.

PR 18-APR-2001; 2001EP-00201407.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Goossens A, Inze D;

PS WPI, 2003-058637/05.

PT Use of expression cassette comprising gene encoding ATP-binding cassette-

PT transporter to induce or enhance production or secretion of secondary

PT metabolite by plant or plant cells.

PS Example 3; Page 49-55; 58pp; English.

XX This invention describes the novel use of an expression cassette

CC comprising a gene encoding an ATP-binding cassette (ABC)-transporter to

CC induce/enhance production/secretion of secondary metabolite by plant

CC cells, or to stimulate production of the secondary metabolite by plants.

CC The method comprises transforming plant cells or the plant with the

CC expression vector containing the expression cassette, selecting the

CC transformed transgenic plant or plant cells, and propagating them. The

CC expression cassette is useful for inducing or enhancing production or

CC secretion of secondary metabolites in plants or plant cells, which e.g.

CC play a role in the resistance of plants to pests and diseases and which

CC can have anti-oxidative and anti-metastatic properties. The secondary

CC metabolites are preferably alkaloids e.g. terpenoid indole alkaloids,

CC tropone alkaloids, steroid alkaloids and polyhydroxy alkaloids. The

CC cassette of the invention has cytostatic and pesticidal activity. This

CC sequence represents the vector pK7MG2D which is used in the expression of

CC the pleiotropic drug resistance genes described in the disclosure of the

CC invention

XX Sequence 12789 BP; 3050 A; 3326 C; 3397 G; 3015 T; 0 U; 1 Other;

QY Query Match 84.7%; Score 720; DB 7; Length 12789;

Best Local Similarity 100.0%; Pred. No. 4.2e-116;

Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTAGAGCAAGGGGAGGAGGCTGTTCAACGGGGTGGTGGCCATCTGCTGAGCTGAGAC 60

DB 5797 ATGGTAGAGCAAGGGGAGGAGGCTGTTCAACGGGGTGGTGGCCATCTGCTGAGCTGAGAC 5738

QY 61 GGCGACGTAACGGGCAACAGTTGAGCGTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 120
 DB 5737 GGCGACGTAACGGGCAACAGTTGAGCGTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 5678
 QY 121 GGCAAGCTGACCCCTGAGTTGATCTGACCAACCGCGAAGCTGCCCTGTCCCTGCGCCACC 180
 DB 5677 GGCAAGCTGACCCCTGAGTTGATCTGACCAACCGCGAAGCTGCCCTGTCCCTGCGCCACC 5618
 QY 181 CTCTGACCACTTGAAGCTGAGCGTGTGAGGCTTCAAGCGGTACCCCGACCAATGAG 240
 DB 5617 CTCTGACCACTTGAAGCTGAGCGTGTGAGGCTTCAAGCGGTACCCCGACCAATGAG 5558
 QY 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCCGAGGCTTCAAGGAGCGGACCAATCTTC 300
 DB 5557 CAGACGACTTCTTCAAGTCCGCGCATGCGCCGAGGCTTCAAGGAGCGGACCAATCTTC 5498
 QY 301 TTCAAGGACGACGGCACTTACAAAGACCGCGCGGAGGTGAGGTTGAGGGCGACACCTTG 360
 DB 5497 TTCAAGGACGACGGCACTTACAAAGACCGCGCGGAGGTGAGGTTGAGGGCGACACCTTG 5438
 QY 361 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGGAGGAGCGGACCAATCTGCGGAC 420
 DB 5437 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGGAGGAGCGGACCAATCTGCGGAC 5378
 QY 421 AAGCTGAGTACAACTTACAAAGCAACGCTTATCATGCGCGACAGCAAGCAAGAAC 480
 DB 5377 AAGCTGAGTACAACTTACAAAGCAACGCTTATCATGCGCGACAGCAAGCAAGAAC 5318
 QY 481 GGCATCAAGTGAAGTTCAGATCCGCGCAACATCGAGAGACGCGATGAGCTCGCC 540
 DB 5317 GGCATCAAGTGAAGTTCAGATCCGCGCAACATCGAGAGACGCGATGAGCTCGCC 5258
 QY 541 GACCACTACCAAGCAACACCGCCCATGCGCGAGCGCGCGCTGCTGCGCGACCAACAC 600
 DB 5257 GACCACTACCAAGCAACACCGCCCATGCGCGAGCGCGCGCTGCTGCGCGACCAACAC 5198
 QY 601 TACCTGAGACCCAGTCCGCTGAGCAAGACCCCAACGAGAAAGCGCGATCAATGTC 660
 DB 5197 TACCTGAGACCCAGTCCGCTGAGCAAGACCCCAACGAGAAAGCGCGATCAATGTC 5138
 QY 661 CTGCTGAGTGTGAGCGCGCGCGGAGTCACTCTGCGGATGAGAGCTGTACAAAG 720
 DB 5137 CTGCTGAGTGTGAGCGCGCGCGGAGTCACTCTGCGGATGAGAGCTGTACAAAG 5078
 RESULT 15
 ID ADB16319 standard; DNA; 7326 BP.
 XX ADB16319;
 AC ADB16319;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Total plasmid pONT8.4TsynCOG/ pONT8.4TsynCOG1 DNA sequence.
 XX
 KM multicistronic; retroviral vector; lentiviral; tetracycline repressor;
 KM TetR; inflammatory disease; asthma; psoriasis; cancer;
 KM chronic myeloid leukemia; CML; neurodegenerative; Alzheimer's; viral; ds;
 KM neotropic; neuroprotective; immunosuppressive; antiinflammatory; ds;
 KM cyclic; circular; pONT8.4TsynCOG; pONT8.4TsynCOG1.
 XX
 OS Unidentified.
 OS Equine infectious anemia virus.
 OS Cytomegalovirus.
 OS
 XX MO2003064665-A2.
 XX
 XX 07-AUG-2003.
 PD
 XX
 PF 03-FEB-2003; 2003MO-GB000418.
 XX
 PR 01-FEB-2002; 2002GB-00002403.
 PR 31-MAY-2002; 2002GB-00012768.
 PR

XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Radcliffe P, Miskin JE, Wilkes FU, Mitrophanous KA;
 XX WPI; 2003-646154/61.
 DR
 XX
 PT New multicistronic retroviral vector genome comprising a first nucleic
 PT acid sequence upstream of an integral regulatory element, useful for
 PT treating inflammation, asthma, viral diseases, psoriasis, or
 PT neurodegenerative diseases.
 XX
 PS Disclosure; Fig 18; 193pp; English.
 XX
 CC This invention relates to a novel multicistronic retroviral vector genome
 CC comprising a first nucleic acid sequence upstream of an integral
 CC regulatory element, such that the level of genomic RNA available for
 CC packaging in the absence of rev, or its functional equivalent, is
 CC increased. Specifically, the vector is a lentiviral genome that comprises
 CC a therapeutically important nucleotide sequence of interest operably
 CC linked to a tetracycline operator. Furthermore, the first nucleic acid
 CC sequence represents a selectable marker such as the tetracycline
 CC repressor (TetR) gene, which is codon optimised for expression in
 CC mammalian cells. Together, the retroviral vector particle, DNA construct
 CC and mammalian cell are useful for the preparation of a medicament to
 CC deliver the therapeutic sequence of interest to a target site. As such,
 CC the present invention describes a method that can be used in the
 CC treatment of inflammatory disease, asthma, psoriasis, cancer e.g. chronic
 CC myeloid leukaemia (CML) and neurodegenerative disease e.g. Alzheimer's.
 CC Accordingly, via gene therapy these vectors can be described as having
 CC antiviral, neurotropic, neuroprotective, immunosuppressive and/or
 CC antiinflammatory activity. This polynucleotide sequence represents the
 CC total plasmid DNA of pONT8.4TsynCOG/ pONT8.4TsynCOG1 of the invention.
 XX
 SQ Sequence 7326 BP; 1880 A; 1853 C; 1843 G; 1750 T; 0 U; 0 Other;
 Query Match 84.6%; Score 719.4; DB 8; Length 7326;
 Best Local Similarity 97.9%; Freq. No. 5.4e-116;
 Matches 729; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 1 ATGTGAGCAAGGGGAGGAGGAGTGTTCACCGGGGTGAGGCCATCTGTGAGCTGAGC 60
 DB 3333 ATGTGAGCAAGGGGAGGAGGAGTGTTCACCGGGGTGAGGCCATCTGTGAGCTGAGC 3392
 QY 61 GGCGACGTAACGGGCAACAGTTGAGCGTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 120
 DB 3393 GGCGACGTAACGGGCAACAGTTGAGCGTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 3452
 QY 121 GGCAAGCTGACCCCTGAGTTGATCTGACCAACCGCGAAGCTGCCCTGTCCCTGCGCCACC 180
 DB 3453 GGCAAGCTGACCCCTGAGTTGATCTGACCAACCGCGAAGCTGCCCTGTCCCTGCGCCACC 3512
 QY 181 CTCTGACCACTTGAAGCTGAGCGTGTGAGGCTTCAAGCGGTACCCCGACCAATGAG 240
 DB 3513 CTCTGACCACTTGAAGCTGAGCGTGTGAGGCTTCAAGCGGTACCCCGACCAATGAG 3572
 QY 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCCGAGGCTTCAAGGAGCGGACCAATCTTC 300
 DB 3573 CAGACGACTTCTTCAAGTCCGCGCATGCGCCGAGGCTTCAAGGAGCGGACCAATCTTC 3632
 QY 301 TTCAAGGACGACGGCACTTACAAAGACCGCGCGGAGGTGAGGTTGAGGGCGACACCTTG 360
 DB 3633 TTCAAGGACGACGGCACTTACAAAGACCGCGCGGAGGTGAGGTTGAGGGCGACACCTTG 3692
 QY 361 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGGAGGAGCGGACCAATCTGCGGAC 420
 DB 3693 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGGAGGAGCGGACCAATCTGCGGAC 3752
 QY 421 AAGCTGAGTACAACTTACAAAGCAACGCTTATCATGCGCGACAGCAAGCAAGAAC 480
 DB 3753 AAGCTGAGTACAACTTACAAAGCAACGCTTATCATGCGCGACAGCAAGCAAGAAC 3812
 QY 481 GGCATCAAGTGAAGTTCAGATCCGCGCAACATCGAGAGACGCGATGAGCTCGCC 540

```
Db      ||||| 3813 GGCATCAAGTGAACCTCAAGATCCGCCACAACATCGAGACGCGAGCTCGCC 3872
QY      ||||| 541 GACCACTACCAAGACAACCCCATCGGCGAGGCGCCGCTGCTGCGCCGACAACAC 600
Db      ||||| 3873 GACCACTACCAAGACAACCCCATCGGCGAGGCGCCGCTGCTGCGCCGACAACAC 3932
QY      ||||| 601 TACCTGAGCACCAGTCCGCCCTGAGCAAGAAGACCCCAACGAGAGCGGATCACAATGATC 660
Db      ||||| 3933 TACCTGAGCACCAGTCCGCCCTGAGCAAGAAGACCCCAACGAGAGCGGATCACAATGATC 3992
QY      ||||| 661 CTGCTGAGAGTTGTGACGCGCGCCGCGGATCACTCTGCGCATGAGAGCTGTACAAGAAG 720
Db      ||||| 3993 CTGCTGAGAGTTGTGACGCGCGCCGCGGATCACTCTGCGCATGAGAGCTGTACAAGAAG 4052
QY      ||||| 721 CTTAGCATGCGCTCCCGCGGAGG 745
Db      ||||| 4053 AGCGGCGCGACTCTAGCTTGCAAG 4077
```

Search completed: May 15, 2004, 03:30:28
Job time : 440 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 02:53:31 ; Search time 2816 Seconds

(without alignments)
9013.797 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atggtgagcgaagcgagga.....tagatcaatgtagatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0 ; Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthm:*
3: em_esthm:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pln:*
20: em_gss_vrl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	515.2	60.6	697	13	BU626888 UI-H-PTO-
2	493.8	58.1	790	13	AL044652 DXFP434P
3	456.4	53.7	635	14	CD369111 UI-H-PTI-
4	438.4	51.6	615	12	BM209799 C0653E03-

Result No.	Score	Query Match	Length	ID	Description
5	417.4	49.1	536	14	CB529588 UI-H-FT2-
6	411.4	48.4	550	13	BO551001 H4006G01-
7	274.8	32.3	456	12	BM210789 C0670D12-
8	273	32.1	467	12	BO034019 UI-1-BY0-
9	234.4	27.6	433	14	CA431885 UI-H-FTO-
10	210.4	24.8	397	9	AJ437938
11	205.2	24.1	317	9	AJ437931
12	205	24.1	511	9	AJ437840
13	168	19.8	570	12	BM194666 EST RERE
14	161.8	19.0	240	12	BM688204 TMM102 Hu
15	161	18.9	525	12	BM736033 TMB102 Hu
16	146	17.2	193	9	AJ437933
17	142	16.7	317	13	BO551000 H4006G01-
18	130.6	15.4	506	10	BE784326 602110101
19	129.8	15.3	428	12	BI965250 i435G04.Y
20	129.8	15.3	465	12	BM205695 C0287D12-
21	129.8	15.3	467	10	BE629419 u437903.Y
22	129.8	15.3	473	10	BE631772 u46411.Y
23	129.8	15.3	473	10	BE861518 UI-M-AH1-
24	129.8	15.3	481	12	BM246643 K0741A07-
25	129.8	15.3	487	14	CF981920 ma175d02.
26	129.8	15.3	498	14	CB524049 UI-M-GK0-
27	129.8	15.3	504	12	BM249605 K0836R06-
28	129.8	15.3	518	13	BU605375 mdh90F10.
29	129.8	15.3	519	9	A1848698 UI-M-AH1-
30	129.8	15.3	526	10	AW541064 C0140F03-
31	129.8	15.3	548	14	CF949990 UI-M-HK0-
32	129.8	15.3	550	13	BO930088 AGNCOURT
33	129.8	15.3	555	12	BM054167 i451d10.Y
34	129.8	15.3	562	12	BM863417 602797075
35	129.8	15.3	562	12	BM241829 K0623B12-
36	129.8	15.3	564	13	BO565843 g146112.Y
37	129.8	15.3	566	9	A1429802 mg04b08.x
38	129.8	15.3	568	12	BM241679 K0620E06-
39	129.8	15.3	570	10	BM055221 H1028R05-
40	129.8	15.3	572	10	AW540851 C0137H08-
41	129.8	15.3	586	12	BM082647 H3078F02-
42	129.8	15.3	590	10	BE630086 u415B05.x
43	129.8	15.3	591	12	BM231966 K0316G05-
44	129.8	15.3	594	14	W98488 mg20911.Y1
45	129.8	15.3	609	12	BM241264 K0614E12-

ALIGNMENTS

RESULT 1
BU626888/c
LOCUS
DEFINITION
UI-H-PTO-bhn-d-01-0-UI.s1 NCI CGAP PT0 Homo sapiens CDNA clone
UI-H-PTO-bhn-d-01-0-UI 3', mRNA sequence.
ACCESSION
BU626888
VERSION
BU626888.1 GI:2293103
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 697)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Unpublished (1997)
Tumor Gene Index
JOURNAL
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Robeff-Panella, U of I
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA-Yes.

FEATURES

Location/Qualifiers
 1..697
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UT-H-FT0-Dm-4-01-0-UT"
 /issue_type="Alveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_FTO"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FTO is a cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCATGCGG. The cell line was provided by Gary W. Hunninghake from the University of Iowa."
 TAG TISSUE=Human Lung Alveolar Macrophage
 TAG LIB=UT-H-FT0
 TAG_SEQ=GGCATGCGG"

ORIGIN

Query Match 60.6%; Score 515.2; DB 13; Length 697;
 Best Local Similarity 98.5%; Pred. No. 2.8e-60;
 Matches 520; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 200 AGGGGCTGAGGCTTCAAGCCGCTACCCGACCATATAGAGACGACCTTCTCAAT 259
 697 AGGGGCTGAGGCTTCAAGCCGCTACCCGACCATATAGAGACGACCTTCTCAAT 638
 260 CGGCATGCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 319
 637 CGGCATGCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 578
 320 ACAAGACCGGCGCGGAGTGAAGTTGAGGGCGACACCTGTGTGAACCGCATGAGCTGA 379
 577 ACAAGACCGGCGCGGAGTGAAGTTGAGGGCGACACCTGTGTGAACCGCATGAGCTGA 518
 380 AGGGGCTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 439
 517 AGGGGCTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 458
 440 ACAGCCACAGCTTATATCATGCGGACGACGAGAGAGGACATCAAGTGAATCTTA 499
 457 ACAGCCACAGCTTATATCATGCGGACGACGAGAGAGGACATCAAGTGAATCTTA 398
 500 AGATCCGCAACAATCGAGAGCGGAGCGTGCAGCTTCCGACCACTACAGACAGACA 559
 397 AGATCCGCAACAATCGAGAGCGGAGCGTGCAGCTTCCGACCACTACAGACAGACA 338
 560 CCCCCATGCGGCGCGGCGGCTGCTGCGGACGACGACCACTACAGACCGGCTCGG 619
 337 CCCCCATGCGGCGCGGCGGCTGCTGCGGACGACGACCACTACAGACCGGCTCGG 278
 620 CCGTGAAGAAAGACCCCAACGAGAGCGGACATCAAGTCTCTGTGAGTTGTAACCG 679
 277 CCGTGAAGAAAGACCCCAACGAGAGCGGACATCAAGTCTCTGTGAGTTGTAACCG 218
 680 CCGCGGGATCACTCTCGGATGAGAGAGCTGTACAAAGCTTAGCC 727
 217 CCGCGGGATCACTCTCGGATGAGAGAGCTGTACAAAGCTTAGCC 170

RESULT 2

AL044652
 LOCUS 790 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZp434P092.r1.434 (synonym: htes3) Homo sapiens cDNA clone
 DKFZp434P092.5, mRNA sequence.
 ACCESSION AL044652
 VERSION AL044652.1 GI:5432867
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 790)
 Ansoorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
 Wiemann,S.
 EST (Ansoorge, Benes, et al.)
 Unpublished (1999)
 COMMENT
 TITLE
 JOURNAL
 AUTHORS
 REFERENCE
 MIPs
 Ingolstaedter Landstr.1, D-85764 Neubherg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de;
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFZp434P092) is available at the RZPD in Berlin.
 Please contact the RZPD: Resourcenentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers
 1..790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp434P092"
 /issue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 58.1%; Score 493.8; DB 9; Length 790;
 Best Local Similarity 82.5%; Pred. No. 2.3e-57;
 Matches 578; Conservative 0; Mismatches 122; Indels 1; Gaps 1;
 17 AGGAGCTTTCAACCGGGTGGTCCCATCTGCTGAGCTGAGCGGACGTAACGGCC 76
 1 AGGAGCTTTCAACCGGGTGGTCCCATCTGCTGAGCTGAGCGGACGTAACGGCC 59
 77 ACAAGTTGAGGCTGCGGCGGACGAGGCGAGTGCACCTACGCGCAAGCTGACCTGA 136
 60 ACAAGTTTCTGTGTCAGCGGAGAGGGTGAAGTGAATGACCAATAGGAAGCTCACCTGA 119
 137 AGTTTATCTGACACCGCGGAAAGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTG 196
 120 AATTCTATCTGACACCTGGAAGGCTCCCTGTGTCATGCGCAACACTGTCTACTACTTCA 179
 197 CCTAGGGGTGAGGCTTCAAGCGGCTACCCGACCATATAGAGACGACGACACTTCTCA 256
 180 CCTAGGGGTGAGGCTTCTTCCAGATACCGAGACCAATAGAGAGCATACCTTTTCA 239
 257 AGTCCGCAATGCGGAGGCTACGTCAGAGCGGACGACCATCTTCTTCAAGAGCAGCGCA 316
 240 AGAGCGCATGCGCGAGGCTATGTCAGAGAGAGAACCATCTTTTCAAGATGACGAGA 299
 317 ACTAAGAACCGGCGGCGGAGGAGTTCAGAGGAGGACACCTGTGTGAACGAGATCGAGC 376
 300 ACTAAGAACCGGCGGCTGAAGTCAAGTTCGAAGGAGACACCTGTGTGAATGAGATCGAGC 359
 377 TGAAGGCGATCGACTTCAAGAGGAGCGGACATCTCTGGGCGACAAAGCTGAGTAACT 436
 360 TGAAGGCGATCGACTTCAAGAGGAGTGAACATCTCTGGGCGACAAAGCTGAGTAACT 419

QY	437	ACAACAGCGCAACAACCTCATATTCATAGGCGCGACAGACAGAGAACCGCATCAAGGTGA	496
Db	420	ATTAATCCCAACATGTGTACATCATAGGCGCGAACAACGAAAGATGGCATCAAGGTCACT	479
QY	437	TCGAATTCGCGCCACACATCGAGGACGGCAGGTGTAGTCTCGCGGACCACTATCCAGCAGA	556
Db	480	TCGAATCAGACACACACATGTAGAGATGATTCGTGCAAGTGGCCACCATTTATCAACGA	539
QY	557	ACACCCCATCGGCGACGCGCCCGGTGCTGCTCCGACACACACTACCTGACGACCCAGT	616
Db	540	ACACTCCATGCGCGACCGGCGCTGTGCTCTCTCCAGACACACATTACTGTCCACCCAGT	599
QY	617	CGGCTTGAAGAAAGACCCCAAAGAAAGCGATCACTGTGCTGTGAGATTCTGA	676
Db	600	CTGCGCTGTCTAAACATCCCAACGAAAGAGACACATGCTCTGCTGAGTTGTGA	659
QY	677	CGGCGCGCGGATCACTCTCGGCGATGAGACGACTGTACAG	717
Db	660	CGGCTGCTGGATCAACATGCGCAGTGAAGAGACTGTACAG	700

FEATURES

source

CD369111 635 bp mRNA linear EST 29-MAY-2003
 UI-H-FT1 b1y-m-07-0-UI.s1 NCI CGAP_FTI Homo sapiens CDNA clone
 UI-H-FT1 b1y-m-07-0-UI 3', mRNA sequence.
 CD369111
 CD369111.1 GI:31153201
 EST.
 Homo sapiens (human)
 Homo sapiens
 Sukhayaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 635)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@rockefeller.edu
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgap.html>
 Seq Primer: M13 FORWARD
 PolyA=Yes.

```

location/Qualifiers
1. .635
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-PT1-bjy-m-07-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP P11"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
NCI CGAP P11 is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldi, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT733-pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is

```

ORIGIN

GGCCATGCGG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=01-H-PT1
TAG_SEQ=GGCCATGCGG"

Query Match	53.7%;	Score 456.4;	DB 14;	Length 635;
Best Local Similarity	99.8%;	Pred. No. 2.5e-52;		
Matches 457;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Oy	262	GCATGCCCCAAGSCTAGTTCAGAGGAGCACTCTTTCAAGAGCAAGGCACTAC	321
Db	635	GCATGCCCGAAGGCTAAGTCCAGAGAGCACTCTTTCAAGAGCAAGGCACTAC	576
Oy	322	AAGA CCGCGCCGAGGTGAAGTTGAGGCGCACCTCTGTGAACCGCATGAGCTGAAG	381
Db	575	AAGACCCGCGCCGAGGTGAAGTTGAGAGGCGACACCTCTGTGAACCGCATGAGCTGAAG	516
Oy	382	GGCATCGACTTCAAGAGAGACGGGACAATCCTGGGGGACAAGCTGAGTAACAATCAAC	441
Db	515	GGCATCGACTTCAAGAGAGAGAGGACAATCTCTGGGGACAAGCTGAGTAACAATCAAC	456
Oy	442	AGCCACAACGCTATATCATGCGCCGACAGACGAGAAGAACGGATCAAGSTGAATTTCAAG	501
Db	455	AGCCACAACGCTATATCATGCGCCGACAGACGAGAAGAACGGATCAAGSTGAATTTCAAG	396
Oy	502	ATCCGCCACAAACATCGAGAGACGGGACAGGTGAGTCTGCCGACCACTACACGACAAACC	561
Db	395	ATCCGCCACAAACATCGAGAGACGGGACAGGTGAGTCTGCCGACCACTACACGACAAACC	336
Oy	562	CCCATTCGGCGACGGCCCGTGTCTGTGCCGACAAACAATCATCTGAGACCCAGCTGCCCC	621
Db	335	CCCATTCGGCGACGGCCCGTGTGTCTGTGCCGACAAACAATCATCTGAGACCCAGCTGCCCC	276
Oy	622	CTGAGCAAAAGACCCCAACGAGAACGCGATCAATGCTCTGTGGAATTGTGACCGCC	681
Db	275	CTGAGCAAAAGACCCCAACGAGAACGCGATCAATGCTCTGTGGAATTGTGACCGCC	216
Oy	682	GGCGGGAATCATCTCGGAGTAGGAGAGAGCTGACAAAGAA	719
Db	215	GGCGGGAATCATCTCGGAGTAGGAGAGAGCTGACAAAGTA	178

RESULT 4

LOCUS	BM209799	615 bp	mRNA	EST	08-JUN-2003
DEFINITION	C0653E03-3 NIA Mouse Trophoblast		Stem Cell	Library	Musculus
DESCRIPTION	musculus cDNA clone C0653E03 3',		mRNA sequence.		

ACCESSION	BM209799	
VERSION	BM209799.2	GI:31533410
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Mus;
1 (bases 1 to 615)
Piao, S., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PubMed 11544199
COMMENT On Dec 14, 2001 this sequence version replaced gi:17766404.

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@nlsun.grc.nia.nih.gov
Plate: C0653 row: E column: 03
Seq primer: -10m3 Forward
High quality sequence stop: 615

FEATURES
source

POLYA=Yes.

Location/Qualifiers

1. 615
/organism="Mus musculus"
/mol_type="mRNA"
/strain="B5/EGFP transgenic ICR mice"
/db_xref="taxon:10090"
/clone="C0653E03"
/issue_type="Trophoblast stem cell"
/dev_stage="3.5-dpc"
/lab_host="DH10B"
/clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library (Short)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a short-transcript enriched cDNA library (Ref. Genome Res. 1: 1553-1558 (2001). [PMID: 11541991]. Total RNAs were obtained from Dr. Janet Rossant and Tilo Kunath (Samuel Lunenfeld Research Institute, Canada). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCCGCTTTT-3'] from 4 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L-SalI, purified by phenol/chloroform, and separated from free linkers by centrifugation 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-L. The products were purified by phenol/chloroform and centrifuged 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.3 Kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 51.6%; Score 438.4; DB 12; Length 615;
Best Local Similarity 86.4%; Pred. No. 6.6e-50;
Matches 484; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

279 CGTCCAGAGCGGACCATCTTTCTTCAAGAGCGGCACTACAGACCGCGCCGAGGT 338
615 CGTCCAGAGCGGACCATCTTTCTTCAAGAGCGGCACTACAGACCGCGCCGAGGT 556
339 GAAGTTCGAGGGGAGCAACCTGTGAAACCGCATCGAGCTGAAGGGCATGTGACTTCAAG 398
555 GAAGTTCGAGGGGAGCAACCTGTGAAACCGCATCGAGCTGAAGGGCATGTGACTTCAAG 496
399 GGAAGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAAGGCTATAT 458
495 GGAAGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAAGGCTATAT 436
459 CATGGCCGACCAAGAGGAGGATCAAGGTGAAGTTCAGATCGCCACACATCGA 518
435 CATGGCCGACCAAGAGGAGGATCAAGGTGAAGTTCAGATCGCCACACATCGA 376
519 GGAAGGCAAGGTGACCTCGCCGACCACTACAGAGGAAACCCCATGCGGACGAGCC 578
375 GGAAGGCAAGGTGACCTCGCCGACCACTACAGAGGAAACCCCATGCGGACGAGCC 316
579 CGTGTGCTGCGCCGACCACTACAGAGGAAACCCCATGCGGACGAGCC 638
315 CGTGTGCTGCGCCGACCACTACAGAGGAAACCCCATGCGGACGAGCC 256
639 CGAAGACCGGATCATGTGCTGTGAGATTGTGACCGCGCGGAGATCATCTTCG 698
255 CGAAGACCGGATCATGTGCTGTGAGATTGTGACCGCGCGGAGATCATCTTCG 196
699 CATGACGAGCTGTACAGAGAGCTTAGCCATGCTCCGCGGAGAGTGAAGAGAGGA 758

DB 195 CATGACGAGCTGTACAGTAGAATTCATCTCTCCTCAGGTGACGGTGCCTATCAAGAGT 136
QY 759 TGAATGACGCTGCCCATGTCTGTGCGCCAGGAGAGCGGAGATGAGACCGTCAACCCGAGC 818
DB 135 GGTGTGCTGTGTGGCAATATGCTGCTGCTCAAAATACATGAGATCTTTTCCCTTC 76
QY 819 CTGTGCTGTGCTGAGATCA 838
DB 75 CAAAATATATGGGACATCA 56

RESULT 5
CB529588/c 596 bp mRNA linear EST 16-MAY-2003
LOCUS
DEFINITION
UI-H-FT2-bjh-d-17-0-UI-s1 NCI CGAP_Ft2 Homo sapiens cDNA clone
UI-H-FT2-bjh-d-17-0-UI 3', mRNA sequence.
CB529588
ACCESSION
CB529588
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 596)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.owa.edu/distribution/cgap.html
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1. 596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bjh-d-17-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_Ft2"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_Ft2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was subtracted according to Bonaldi, Lemon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FT2
TAG_SEQ=GCCATGCG"

ORIGIN

Query Match 49.1%; Score 417.4; DB 14; Length 596;
Best Local Similarity 99.8%; Pred. No. 4.5e-47;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 301 TTCAAGAGAGAGGCACTACAGACCGCGCCGAGGTGAATTCAGAGGCGGACCCCTG 360
DB 596 TTCAAGAGAGAGGCACTACAGACCGCGCCGAGGTGAATTCAGAGGCGGACCCCTG 537
QY 361 GTGAACCGGATGAGCTGAGGGGATGACCTTCAAGAGAGAGCGGACACATCTCTGGGAC 420
DB 536 GTGAACCGGATGAGCTGAGGGGATGACCTTCAAGAGAGAGCGGACACATCTCTGGGAC 477

Qy 421 AAGCTGAGTACAACTACAAAGCAAGCAACGTATATCATGCGGACAGCAAGAAC 480
Db 476 AAGCTGAGTACAACTACAAAGCAAGCAACGTATATCATGCGGACAGCAAGAAC 417
Qy 481 GGCATCAAGGTGAATTCAGATCCGCCCAACATCGAGAGCGGACGCTGCACTGCC 540
Db 416 GGCATCAAGGTGAATTCAGATCCGCCCAACATCGAGAGCGGACGCTGCACTGCC 357
Qy 541 GACCACTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 600
Db 356 GACCACTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 297
Qy 601 TACCTGACACCACTGACCTGACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCA 660
Db 296 TACCTGACACCACTGACCTGACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCA 237
Qy 661 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTGCGCATGAGCAGCTGTACAGAA 719
Db 236 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTGCGCATGAGCAGCTGTACAGAA 178

RESULT 6
BOS51001 550 bp mRNA linear EST 20-JUN-2002
LOCUS H4006G01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
H4006G01 5', mRNA sequence.
BOS51001
ACCESSION BOS51001.1 GI:21451887
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS VanBuren V., Piao Y., Dudekula D.B., Qian Y., Carter M.G.,
Martin P.R., Stagg C.A., Bassey U., Alpa K., Hamatani T.,
Kargul G.J., Luo A.G., Kelso J., Hide W. and Ko M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
JOURNAL Genome Res. 12 (12), 1999-2003 (2002)
MEDLINE 22354164
PUBMED 12466305

COMMENT
Other ESTs: H400G01-3
Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cda@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
Plate: H4006 row: G column: 01
Seq primer: -21M13 Reverse
High quality sequence stop: 550
POLYA=No.

FEATURES
Source
1. .550
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="H400G01"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

ORIGIN
Query Match 48.4%; Score 411.4; DB 13; Length 550;

Best Local Similarity 99.8%; Pred. No. 2.9e-46;
Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGTGAGCAAGGCGGAGAGCTTTCAACGGGGTGTGCTCCATCTGCTGAGCTGAC 60
Db 138 ATGTGAGCAAGGCGGAGAGCTTTCAACGGGGTGTGCTCCATCTGCTGAGCTGAC 197
Qy 61 GGCAGCTAAACGGCGCAAGTTGAGGCTGTCGGGCGAGGGCGAGGGCGAATGCCACTAC 120
Db 198 GGCAGCTAAACGGCGCAAGTTGAGGCTGTCGGGCGAGGGCGAGGGCGAATGCCACTAC 257
Qy 121 GGCAGCTAAACGGCGCAAGTTGAGGCTGTCGGGCGAGGGCGAGGGCGAATGCCACTAC 180
Db 258 GGCAGCTAAACGGCGCAAGTTGAGGCTGTCGGGCGAGGGCGAGGGCGAATGCCACTAC 317
Qy 181 CTGCTGACCACTGACCTTACGCGGTGAGTGTCTTACGCGGTACCCGACCAATGAG 240
Db 318 CTGCTGACCACTGACCTTACGCGGTGAGTGTCTTACGCGGTACCCGACCAATGAG 377
Qy 241 CAGCAGCACTTCTTCAAGTCGCGGATGCGGAGGCTACGTCAGAGGCGGACCATCTTC 300
Db 378 CAGCAGCACTTCTTCAAGTCGCGGATGCGGAGGCTACGTCAGAGGCGGACCATCTTC 437
Qy 301 TTCAAGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGCGGACCATCTTC 360
Db 438 TTCAAGACGACGCGCAACTACAGACCCGCGCGGAGTGAAGTTGAGGCGGACCATCTTC 497
Qy 361 GTGAACCGCATCGAGCTGAAGGGGATGACTTCAAGAGAGCGGACCATCTTC 413
Db 498 GTGAACCGCATCGAGCTGAAGGGGATGACTTCAAGAGAGCGGACCATCTTC 550

RESULT 7
BMS210789/c 456 bp mRNA linear EST 08-JUN-2003
LOCUS C0670D12-3 NIA Mouse Trophoblast Stem Cell cDNA Library (Short) Mus
musculus cDNA clone C0670D12 3', mRNA sequence.
BMS210789
ACCESSION BMS210789.2 GI:31534239
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS VanBuren V., Piao Y., Dudekula D.B., Qian Y., Carter M.G.,
Martin P.R., Stagg C.A., Bassey U., Alpa K., Hamatani T.,
Kargul G.J., Luo A.G., Kelso J., Hide W. and Ko M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199

COMMENT
On Dec 14, 2001 this sequence version replaced gi:17767758.
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cda@igsun.grc.nia.nih.gov
Plate: C0670 row: D column: 12
Seq primer: -21M13 Forward
High quality sequence stop: 456
POLYA=Yes.

FEATURES
Source
1. .456
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="B6/EGFP transgenic ICR mice"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="C0670D12"
/tissue_type="Trophoblast stem cell"
/dev_stage="3.5-dpc"
/lab_host="DH10B"

/clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library (short)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://19sun.grc.nia.nih.gov/cDNA>). This is a short-transcript enriched cDNA library (Ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Janet Rossant and Tilo Kunath (Samuel Lunenfeld Research Institute, Canada). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTTATGATGCGAGCGCGCCCTTTTCTTTT-3'] from 4 µg of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-L. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.3 Kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 32.3%; Score 274.8; DB 12; Length 456;
 Best Local Similarity 80.7%; Pred. No. 8e-28;
 Matches 321; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

444 CAGCCCAACGTTATATCATGCGCCGACAGAGAGAGCGCATCAAGTGAATTCAA 500
 456 GAGCCCAACGTTATATCATGCGCCGACAGAGAGAGCGCATCAAGTGAATTCAA 397
 501 GATCGCCCAACATGAGAGAGCGCGCTGCGCCGACCTTACGACAGAACAC 560
 396 GATCGCCCAACATGAGAGAGCGCGCTGCGCCGACCTTACGACAGAACAC 337
 561 CCCCATCGGCGAGCGCGCGCTGCTGCGCCGACCTTACGACAGAACAC 620
 336 CCCCATCGGCGAGCGCGCGCTGCTGCGCCGACCTTACGACAGAACAC 277
 621 CCTGAGCAAAAGACCCCAAGAGAGCGCATCATGCTGCTGAGTTCGTGACCCG 680
 276 CCTGAGCAAAAGACCCCAAGAGAGCGCATCATGCTGCTGAGTTCGTGACCCG 217
 681 CGCGGGGATCATCTCTGCGATGAGAGAGCTGTACAAAGAGCTTACGAGGTGCA 740
 216 CGCGGGGATCATCTCTGCGATGAGAGAGCTGTACAAAGAGCTTACGAGGTGCA 157
 741 GAGAGTGAAGAGAGAGATGATGAGAGCGCTGCTGCTGAGGAGAGAGCGGANT 800
 156 GAGTGCCTTACAGAAAGTGTGCTGTGTGAGGAGAGCGCTGCTGAGTACCAAT 97
 801 GAGAGTGAAGAGAGATGATGAGAGCGCTGCTGCTGAGGAGAGAGCGGANT 838
 96 AGATCTTTTCCCTGCGCAAAATATATGAGGAGATCA 59

RESULT 8
 B0034019/c 467 bp mRNA linear EST 27-MAR-2002
 LOCUS UI-1-BY0-ahv-c-10-0-UI.s1 NCI CGAP Tr1 Mus musculus cDNA clone
 DEFINITION UI-1-BY0-ahv-c-10-0-UI 3', mRNA sequence.
 ACCESSION B0034019
 VERSION B0034019.1 GI:19769298
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 467)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Dr. Janet Rossant
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@iowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers

1..467
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-1-BY0-ahv-c-10-0-UI"
 /tissue_type="Trophoblast"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Tr1"
 /note="Organ: Trophoblast; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Tr1 is a cDNA library containing the following tissue(s): Placenta Trophoblast. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGGTG. TAG TISSUE=mouse trophoblast
 TAG LIB=UI-1-BY0
 TAG_SEQ=GGTTG"

ORIGIN

Query Match 32.1%; Score 273; DB 12; Length 467;
 Best Local Similarity 80.5%; Pred. No. 1.4e-27;
 Matches 318; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

444 CCACAACGTTATATCATGCGCCGACAGAGAGAGCGCATCAAGTGAATTGAAGT 503
 467 CCACAACGTTATATCATGCGCCGACAGAGAGAGCGCATCAAGTGAATTGAAGT 408
 504 CGCGCAAAACATGAGAGAGCGAGAGTGTGCGCCGACCTACAGAGAACACCCG 563
 407 CGCGCAAAACATGAGAGAGCGAGAGTGTGCGCCGACCTACAGAGAACACCCG 348
 564 CATCGCGCGAGCGCGCGCTGCTGCGCCGACCTACAGAGAACACCCGCTT 623
 347 CATCGCGCGAGCGCGCGCTGCTGCGCCGACCTACAGAGAACACCCGCTT 288
 624 GAGCAAAAGACCCCAAGAGAGCGAGATCATGCTGCTGAGGAGAGAGCGCGG 683
 287 GAGCAAAAGACCCCAAGAGAGCGAGATCATGCTGCTGAGGAGAGAGCGCGG 228
 684 CGGATCACTCTCGGATGAGAGAGCTGTACAAAGAGCTTACGAGTTCGCCGCGGA 743
 227 CGGATCACTCTCGGATGAGAGAGCTGTACAAAGAGCTTACGAGTTCGCCGCGG 168
 744 GGTGAGAGAGAGAGATGATGAGAGCGTGCCTGCTGCTGAGGAGAGAGCGGATGGA 803
 167 TGCCTATCAGAAAGTGTGCTGTGTGAGGAGAGCGCTGCTGAGTACCAATTA 108
 804 CGGTACCTGAGAGCTGCTGCTGAGTGAATCA 838

Db 107 TCTTTTCCCTCTGCGCAAAATTTATGGGACATCA 73

RESULT 9

CA431885/c

LOCUS CA431885 433 bp mRNA linear EST 07-NOV-2002

DEFINITION

UI-H-F70-bhl-m-13-0-UI-s1 NCI CGAP F70 Homo sapiens cDNA clone

ACCESSION

CA431885

VERSION

CA431885.1 GI:24794611

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 433)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

FEATURES

Source

LOCATION/Qualifiers

1..433

ORGANISM

Homo sapiens

mol_type

"mRNA"

db_xref

"taxon:9606"

clone

"UI-H-F70-bhl-m-13-0-UI"

tissue

"Alveolar Macrophage"

dev stage

"Adult"

lab host

"DH10B (Life Technologies)"

clone_lib

"NCI-CGAP F70"

note

"Organ: Lung; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP F70 is a cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tag for this library is GGCATGCGG. The cell line was provided by Gary W. Hunninghake from the University of Iowa.

TAG

TAG_LIB=UI-H-F70

TAG_SEQ

GGCCATGCGG

ORIGIN

Query Match

Best Local Similarity 27.6%; Score 234.4; DB 14; Length 433;

Matches

249; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY

465 CGACAAGCAGAGAAAGCGCATCAAGTGAATTCAAGAT-CCGCCACAACATCGAGAGACG 523

DB

433 CGACAAGCAGAGAAAGCGCATCAAGTGAATTCAAGAT-CCGCCACAACATCGAGAGAG 374

QY

524 GGAAGGTGAGCTCGCCGACCACTACAGAGAAACCCCGCATGCGCGAGCCCGGTCG 583

DB

373 GGAAGGTGAGCTCGCCGACCACTACAGAGAAACCCCGCATGCGCGAGCCCGGTCG 314

QY 584 TGCTGCCGACCAACACTACTGAGACCCAGTCCGCTTGAGCAAAAGACCCCAAGAGA 643

DB

313 TGCTGCCGACCAACACTACTGAGACCCAGTCCGCTTGAGCAAAAGACCCCAAGAGA 254

QY

644 AGCCGATCAGATGCTCTGCTGAGTTCTGAGACCCGCGCGGATCACTCGGCAATGG 703

DB

253 AGCCGATCAGATGCTCTGCTGAGTTCTGAGACCCGCGCGGATCACTCGGCAATGG 194

QY

704 ACAGCTGTACAGAA 719

DB

193 ACAGCTGTACAGAA 178

RESULT 10

At437938

LOCUS

At437938 Hapd1R2 Helianthus annuus cDNA clone Hapd1R259C04, mRNA

DEFINITION

sequence.

ACCESSION

At437938

VERSION

At437938.1 GI:19525477

KEYWORDS

EST.

SOURCE

Helianthus annuus (common sunflower)

ORGANISM

Helianthus annuus

Eukaryota

Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE

1 (bases 1 to 397)

AUTHORS

Tamborindéguy, C.

TITLE

An expressed-sequence-tags database of the sunflower protoplast

COMMENT

Thesis (2002) Department of Ecole Nationale Supérieure Agronomique de Toulouse, Institut National Polytechnique de Toulouse, Toulouse, France

UNPUBLISHED (2002)

Contact: Genzibittel, L

LABORATOIRE

de Biotechnologie et Amélioration des Plantes

INSTITUT

National Polytechnique de Toulouse - Ecole Nationale Supérieure Agronomique de Toulouse

IFR40

Pôle de Biotechnologie Végétale, 18 chemin de Borde Rouge, Auzeville, CASTANET TOLOSAN 31326, France.

LOCATION/Qualifiers

1..397

ORGANISM

"Helianthus annuus"

mol_type

"mRNA"

db_xref

"taxon:4232"

clone

"Hapd1R259C04"

tissue

"hypocotyls"

cell type

"protoplasts"

dev stage

"1-5 days"

clone_lib

"Hapd1R2"

ORIGIN

Query Match 24.8%; Score 210.4; DB 9; Length 397;

Best Local Similarity

99.5%; Pred. No. 3.9e-19;

Matches

211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 ATGCTGAGCAGAGGAGAGAGCTGTTACCGGGGTGAGCCATCTGATCGAGCTGAGC 60

DB

148 ATGCTGAGCAGAGGAGAGAGCTGTTACCGGGGTGAGCCATCTGATCGAGCTGAGC 207

QY

61 GCGCAGTAAAGCGCAGCAAGTTAGCGTGTCCGGCAGAGGCGAGGCGAGTCCACTTAC 120

DB

208 GCGCAGTAAAGCGCAGCAAGTTAGCGTGTCCGGCAGAGGCGAGGCGAGTCCACTTAC 267

QY

121 GCGCAGTAAAGCGCAGCAAGTTAGCGTGTCCGGCAGAGGCGAGGCGAGTCCACTTAC 180

DB

268 GCGCAGTAAAGCGCAGCAAGTTAGCGTGTCCGGCAGAGGCGAGGCGAGTCCACTTAC 327

QY

181 CTCGTGACCAACCTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 212

DB

328 CTCGTGACCAACCTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 359

RESULT 11
AJ437931 317 bp mRNA linear EST 15-MAR-2002
LOCUS AJ437931
DEFINITION Helianthus annuus cDNA clone HADpLR259A02, mRNA
ACCESSION AJ437931
VERSION AJ437931.1 GI:19525470
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 317)
Tamborindeguy, C.
An expressed-sequenced-tags database of the sunflower protoplast
Thesis (2002) Department of Ecole Nationale Supérieure Agronomique
de Toulouse, Institut National Polytechnique de Toulouse, Toulouse,
France
JOURNAL
Contact: Gentrabitel L
Laboratoire de Biotechnologie et Amélioration des Plantes
Institut National Polytechnique de Toulouse - Ecole National
Supérieure Agronomique de Toulouse
IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
Auzeville, CASTANET TOLOSAN 31326, France.
FEATURES
source
1. 317
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="Emil"
/db_xref="taxon:4232"
/clone="HADpLR259A02"
/tissue_type="hypocotyls"
/cell_type="protoplasts"
/dev_stage="1-5 days"
/clone_1lb="HADpLR2"
ORIGIN
Query Match 24.1%; Score 205.2; DB 9; Length 317;
Best Local Similarity 98.6%; Pred. No. 1.9e-18;
Matches 207; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGNGTCCCATCTGGTGCAGCTGAGC 60
Db 68 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGNGTCCCATCTGGTGCAGCTGAGC 127
QY 61 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 120
Db 128 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 187
QY 121 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 180
Db 188 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 247
QY 181 CTCTGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 210
Db 248 CTCTGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 277

RESULT 12
AJ437840 511 bp mRNA linear EST 15-MAR-2002
LOCUS AJ437840
DEFINITION Helianthus annuus cDNA clone HADpLR205H12, mRNA
ACCESSION AJ437840
VERSION AJ437840.1 GI:19525379
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 511)
Tamborindeguy, C.
An expressed-sequenced-tags database of the sunflower protoplast
Thesis (2002) Department of Ecole Nationale Supérieure Agronomique
de Toulouse, Institut National Polytechnique de Toulouse, Toulouse,
France
JOURNAL
Contact: Gentrabitel L
Laboratoire de Biotechnologie et Amélioration des Plantes
Institut National Polytechnique de Toulouse - Ecole National
Supérieure Agronomique de Toulouse
IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
Auzeville, CASTANET TOLOSAN 31326, France.
FEATURES
source
1. 511
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="Emil"
/db_xref="taxon:4232"
/clone="HADpLR205H12"
/tissue_type="hypocotyls"
/cell_type="protoplasts"
/dev_stage="1-5 days"
/clone_1lb="HADpLR2"
ORIGIN
Query Match 24.1%; Score 205; DB 9; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.2e-18;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGNGTCCCATCTGGTGCAGCTGAGC 60
Db 262 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGNGTCCCATCTGGTGCAGCTGAGC 321
QY 61 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 120
Db 322 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 381
QY 121 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 180
Db 382 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCCACTTAC 441
QY 181 CTCTGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 205
Db 442 CTCTGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 466

RESULT 13
BM194666 570 bp mRNA linear EST 13-DEC-2002
LOCUS BM194666
DEFINITION EST RRRB sh B3+ SSH Library from Adult Cirrhotic Liver Homo sapiens
cDNA, mRNA sequence.
ACCESSION BM194666
VERSION BM194666.1 GI:2665847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 570)
Shackel, N.A., McGunness, P., Abbott, A., Gorrell, G., and McCaughan, G.
Suppression subtractive hybridisation analysis of intrahepatic
differential gene expression in human liver cirrhosis
Unpublished (2002)
JOURNAL
Contact: Shackel NA
Liver Immunobiology Laboratory
Centenary Institute
Locked Bag No. 6, Newtown, NSW 2042, Sydney, Australia
Tel: 61 2 95656101
Fax: 61 2 95656161

Email: n.shackel@centenary.usyd.edu.au
Seq primer: M13 forward.

FEATURES

Location/Qualifiers
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="SSH library from Adult Cirrhotic Liver"
/note="Organ: Liver; Vector: pGEM T Easy II (Promega)"

ORIGIN

Query Match 19.8%; Score 168; DB 12; Length 570;
Best Local Similarity 99.4%; Pred. No. 2.2e-13;
Matches 179; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGCTGAGCAAGGGCGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCGAGCTGAC 60
DB 391 ATGGTGAGCAAGGGCGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCGAGCTGAC 450
QY 61 GGGCAGCTAAACGGCCCAAGTTGAGCTGTCGGCCGAGGGCGAGGGCGATGCCACTAC 120
DB 451 GGGCAGCTAAACGGCCCAAGTTGAGCTGTCGGCCGAGGGCGAGGGCGATGCCACTAC 510
QY 121 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGTGCCTGCTGCGCCAC 179
DB 511 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGTGCCTGCTGCGCCAC 570

RESULT 14
BM888204 240 bp mRNA linear EST 08-MAR-2002
LOCUS TMW102 Human Trabecular Meshwork cDNA library Homo sapiens cDNA 5',

DEFINITION
mRNA sequence.

ACCESSION
BM888204
VERSION
BM888204.1 GI:19271948

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS
1 (bases 1 to 240)
Wirtz M.K., Samples J.R., Xu H., Severson T. and Acott T.S.

TITLE
Expression Profile and Genome Location of cDNA Clones from an
Infant Human Trabecular Meshwork Library

JOURNAL
Unpublished (2002)

COMMENT
Contact: Wirtz MK

Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA

Tel: 503-494-4698

Fax: 503-494-6875

Email: wirtzm@ohsu.edu

Seq primer: 17 Reverse.

Location/Qualifiers

FEATURES

source

1..240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Human Trabecular Meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOPI0F"
/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
cDNA library made from RNA isolated from Trabecular
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the RNA from the frozen cells using a pCDNA3 vector
and TPO10F host cells."

ORIGIN

Query Match 19.0%; Score 161.8; DB 12; Length 240;

Best Local Similarity 94.6%; Pred. No. 1.3e-12;
Matches 174; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 ATGGTGAGCAAGGGCGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCGAGCTGAC 60
DB 57 ATGGTGAGCAAGGGCGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCGAGCTGAC 116
QY 61 GGGCAGCTAAACGGCCCAAGTTGAGCTGTCGGCCGAGGGCGAGGGCGATGCCACTAC 120
DB 117 GGGCAGCTAAACGGCCCAAGTTGAGCTGTCGGCCGAGGGCGAGGGCGATGCCACTAC 175
QY 121 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGTGCCTGCTGCGCCAC 180
DB 176 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGTGCCTGCTGCGCCAC 225
QY 181 CTCG 184
DB 236 CTAG 239

RESULT 15
BM736033

LOCUS TDSUBC_2D10 T7 Zebrafish shield stage whole embryo Danio rerio cDNA
DEFINITION
clone TDSUBC_2D10, mRNA sequence.

ACCESSION
BM736033
VERSION
BM736033.1 GI:19057366

KEYWORDS
EST.

SOURCE
Danio rerio (zebrafish)

ORGANISM
Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 525)

Dickmeis T., Fischer N. and Straehle U.

Nodal targets identified by subtractive hybridization

Unpublished (2001)

Contact: Dickmeis T, Plessey C.

Uwe Straehle

IGBMC

1, rue Laurent Fries, 67404 Illkirch Cedex, France

Tel: +33388653338

Fax: +33388653201

Email: dickmeis@ictus.u-strasbg.fr

subtracted probe clones, enriched for nodal targets.

Location/Qualifiers

FEATURES

source

1..525
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone_lib="TDSUBC_2D10"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="TOPI0 OneShot cells"
/clone_lib="Zebrafish shield stage whole embryo"
/note="Vector: pCRIT-TOPO, directly subcloned fragments
resulting from the nested PCR of the Clontech PERSselect
subtraction procedure"

ORIGIN

Query Match 18.9%; Score 161; DB 12; Length 525;
Best Local Similarity 75.2%; Pred. No. 1.9e-12;
Matches 200; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 460 ATGGCCGACAGGACAGAGAGCGCATCAAGTGAATTCAGATCCGCGACACATCGAG 519
DB 2 ATGGTGTCTGCGAAGAGAGCGCATCAAGCACTTCAAGACCGCGACACATCGAA 61
QY 520 GACGCGAGGTGAGCTGCGACCACTACACGAGAGACACCCCATAGGCGAGGCGCC 579
DB 62 GACGCGAGGTGAGCTGAGTATATCAACAAATATCTCCATTGGCGATACCTT 121
QY 580 GTGCTGCTGCGCGAAGACACACTGAGCAACCGAGTCCGCTGAGGAGAAAGACCCAC 639

Db	122	GTCTTTACGACACCATTTACCTGTCACACAACTGCGCTTTCGAAAGATCCCAAC	181
Qy	640	GAGAGCCGCGATCATGTCTCTGAGTCTGACCGCCGCGGATCATCTCGGC	699
Db	182	GAAAAGAGAGACCATGTGCTCTTGTGATTGTAAAGCTGTGGGATTACACATGAC	241
Qy	700	ATGAGAGCTGTACAGAAAGCTTAG	725
Db	242	ATGATGAATATACAAATATATCTAG	267

Search completed: May 15, 2004, 05:20:30
 Job time : 2821 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 02:57:45 ; Search time 88 Seconds
(without alignments)
5360.321 Million cell updates/sec

Title: US-09-931-232-2
Perfect score: 850
Sequence: 1 atcggtgacgaagcgagcgagga.....tagatcaatgtagatgc 850

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	850	100.0	US-09-062-102-2	Sequence 2, Appli
2	850	100.0	US-09-364-946-2	Sequence 2, Appli
3	718.6	84.5	US-09-770-315-1	Sequence 1, Appli
4	718	84.5	US-09-503-799-1	Sequence 1, Appli
5	717.4	84.4	US-09-172-063-11	Sequence 11, Appli
6	717.4	84.4	US-09-316-919-12	Sequence 11, Appli
7	717.4	84.4	US-09-602-641-11	Sequence 11, Appli
8	717.4	84.4	US-09-920-922-1	Sequence 1, Appli
9	717.4	84.4	US-09-172-063-29	Sequence 29, Appli
10	717.4	84.4	US-09-602-641-29	Sequence 29, Appli
11	717.4	84.4	US-09-085-305-5	Sequence 5, Appli
12	717.4	84.4	US-09-417-197-128	Sequence 128, App
13	717.4	84.4	US-09-417-197-112	Sequence 112, App
14	717.4	84.4	US-09-417-197-58	Sequence 58, Appli
15	717.4	84.4	US-09-417-197-64	Sequence 64, Appli
16	717.4	84.4	US-09-417-197-62	Sequence 62, Appli
17	717.4	84.4	US-08-818-253-1	Sequence 1, Appli
18	717.4	84.4	US-08-818-253-5	Sequence 5, Appli
19	717.4	84.4	US-08-818-252-1	Sequence 1, Appli
20	717.4	84.4	US-08-818-252-5	Sequence 5, Appli
21	717.4	84.4	US-09-800-170-47	Sequence 47, Appli
22	717.4	84.4	US-09-417-197-74	Sequence 74, Appli
23	717.4	84.4	US-09-417-197-70	Sequence 70, Appli
24	717.4	84.4	US-09-417-197-140	Sequence 140, App
25	717.4	84.4	US-09-417-197-76	Sequence 76, Appli
26	717.4	84.4	US-09-417-197-60	Sequence 60, Appli
27	717.4	84.4	US-09-417-197-118	Sequence 118, App

28	717.4	84.4	US-09-591-025-8	Sequence 8, Appli
29	717.4	84.4	US-09-894-927B-8	Sequence 8, Appli
30	717.4	84.4	US-09-417-197-110	Sequence 110, App
31	717.4	84.4	US-09-417-197-136	Sequence 66, App
32	717.4	84.4	US-09-417-197-66	Sequence 122, App
33	717.4	84.4	US-09-417-197-122	Sequence 78, Appli
34	717.4	84.4	US-09-417-197-78	Sequence 132, App
35	717.4	84.4	US-09-417-197-132	Sequence 4, Appli
36	717.4	84.4	US-09-796-575-4	Sequence 5, Appli
37	717.4	84.4	US-09-208-827-5	Sequence 5, Appli
38	717.4	84.4	US-09-208-827-5	Sequence 5, Appli
39	717.4	84.4	US-10-043-074-5	Sequence 2, Appli
40	717.4	84.4	US-09-133-944-2	Sequence 2, Appli
41	717.4	84.4	US-09-208-827-2	Sequence 2, Appli
42	717.4	84.4	US-10-043-074-2	Sequence 1, Appli
43	717.4	84.4	US-09-133-944-1	Sequence 1, Appli
44	717.4	84.4	US-09-208-827-1	Sequence 1, Appli
45	717.4	84.4	US-10-043-074-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-062-102-2
Sequence 2, Application US/09062102
Patent No. 6130313
GENERAL INFORMATION:
APPLICANT: Li, Xiangfang
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
FILE REFERENCE: D6100
CURRENT APPLICATION NUMBER: US/09/062,102
CURRENT FILING DATE: 1998-04-17
EARLIER APPLICATION NUMBER: US 60/060, 855
EARLIER FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: DNA sequence which encodes the EGFP-MOD422-461
OTHER INFORMATION: fusion protein.
US-09-062-102-2

Query Match 100.0%; Score 850; DB 3; Length 850;
Best Local Similarity 100.0%; Pred. No. 1.2e-160; Indels 0; Gaps 0;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGACGAAGGCGAGAGCTGTTCAACGGGAGTGGCCATCTGTGAGCTGAGC 60
Db 1 ATGCTGACGAAGGCGAGAGCTGTTCAACGGGAGTGGCCATCTGTGAGCTGAGC 60
QY 61 GGCACGTAACGGCCCAAGTTCAGCTGTCCGCGCAGGCGGAGGCGGATGCCACTAC 120
Db 61 GGCACGTAACGGCCCAAGTTCAGCTGTCCGCGCAGGCGGAGGCGGATGCCACTAC 120
QY 121 GCGAAGTGAACCCGAGAGTTCATCTGACCAACCGGAGAGTGCCTGAGCCGAC 180
Db 121 GCGAAGTGAACCCGAGAGTTCATCTGACCAACCGGAGAGTGCCTGAGCCGAC 180
QY 181 CTCGTGACCAACCCGAGAGTTCATCTGACCAACCGGAGAGTGCCTGAGCCGAC 240
Db 181 CTCGTGACCAACCCGAGAGTTCATCTGACCAACCGGAGAGTGCCTGAGCCGAC 240
QY 241 CAGACGACCTTCTTCAAGTCCGATGCGGACCGGAGGCTTCCGAGGCGGACCATCTTC 300
Db 241 CAGACGACCTTCTTCAAGTCCGATGCGGACCGGAGGCTTCCGAGGCGGACCATCTTC 300
QY 301 TTCAGGACGAGGCGGACCATCTTCAAGTCCGATGCGGACCGGAGGCTTCCGAGGCGGACCATCTTC 360
Db 301 TTCAGGACGAGGCGGACCATCTTCAAGTCCGATGCGGACCGGAGGCTTCCGAGGCGGACCATCTTC 360

Db 301 TTCAAGGACGACGGAACCTAACAAGCCGCGCGAGTGAAGTTCAGAGGCGACACCTTG 360
Qy 361 GTGAACCGCATCGAGCTGTAAGGCGATCTTCAAGAGAGAGCGCAATATCTGGGCGAC 420
Db 361 GTGAACCGCATCGAGCTGTAAGGCGATCTTCAAGAGAGAGCGCAATATCTGGGCGAC 420
Qy 421 AAGCTGAGTCAACCTAACAAGCGCAACGCTATCATGCGCGAACAAGAAAC 480
Db 421 AAGCTGAGTCAACCTAACAAGCGCAACGCTATCATGCGCGAACAAGAAAC 480
Qy 481 GGCATCAAGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCTGCAAGTCCGC 540
Db 481 GGCATCAAGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCTGCAAGTCCGC 540
Qy 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Db 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Qy 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCGCAAGAGCGAGTCAATGCTC 660
Db 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCGCAAGAGCGAGTCAATGCTC 660
Qy 661 CTGCTGAGTTCGTAGACCGCGCGGATCACTTCGCGATGAGCAAGTGAACAAG 720
Db 661 CTGCTGAGTTCGTAGACCGCGCGGATCACTTCGCGATGAGCAAGTGAACAAG 720
Qy 721 CTGAGCATGCTTCCTCCCGCGAGGTGAGAGAGAGAGATGAGACGCTGCCATGCT 780
Db 721 CTGAGCATGCTTCCTCCCGCGAGGTGAGAGAGAGAGATGAGACGCTGCCATGCT 780
Qy 781 TGTGCGCAGAGAGAGGAGGATGAGACCGTCACTTCGAGCTGCTTCTAGATCAAT 840
Db 781 TGTGCGCAGAGAGAGGAGGATGAGACCGTCACTTCGAGCTGCTTCTAGATCAAT 840
Qy 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 2
US-09-364-946-2
; Sequence 2, Application US/09364946
; Patent No. 6306600
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; APPLICANT: Li, Xiangyang
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP/D2
; CURRENT APPLICATION NUMBER: US/09/364,946
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/191,233
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 850
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURES:
; OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
; OTHER INFORMATION: fusion protein.
US-09-364-946-2

Query Match 100.0%; Score 850; DB 4; Length 850;
Best Local Similarity 100.0%; Pred. No. 1.2e-160; Indels 0; Gaps 0;
Matches 850; Conservative 0; Mismatches 0

Qy 1 ATGGTGAAGAGGAGGAGAGAGTTCACCGGAGTGTGCTCCATCTGTGAGCTGGAC 60
Db 1 ATGGTGAAGAGGAGGAGAGAGTTCACCGGAGTGTGCTCCATCTGTGAGCTGGAC 60
Qy 61 GCGACGTAAACGCGCACAGTTCAAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 120
Db 61 GCGACGTAAACGCGCACAGTTCAAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 120

Db 61 GCGACGTAAACGCGCACAGTTCAAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 120
Qy 121 GGCAGCTGACCCCTGAGAGTTCATCTGACCAACCGGAAAGTGTCCCTGCTGACCTAC 180
Db 121 GGCAGCTGACCCCTGAGAGTTCATCTGACCAACCGGAAAGTGTCCCTGCTGACCTAC 180
Qy 181 CTGTGACCAACCTGACCTTACAGGCTGACAGTCTTCAAGCGGCTTACCCCGCAACATGAG 240
Db 181 CTGTGACCAACCTGACCTTACAGGCTGACAGTCTTCAAGCGGCTTACCCCGCAACATGAG 240
Qy 241 CAGCAGACCTTCAAGTCCGCGCATCCGAGAGGCTACAGAGGCGCAACCATCTTC 300
Db 241 CAGCAGACCTTCAAGTCCGCGCATCCGAGAGGCTACAGAGGCGCAACCATCTTC 300
Qy 301 TTCAAGGACGACGCAACCTAACAAGCCGCGCGAGGTGAAGTTCAGAGGCGAACCCTG 360
Db 301 TTCAAGGACGACGCAACCTAACAAGCCGCGCGAGGTGAAGTTCAGAGGCGAACCCTG 360
Qy 361 GTGAACCGCATCGAGCTGTAAGGCGATGACTTCAAGAGAGAGCGCAACCTGCGGCGAC 420
Db 361 GTGAACCGCATCGAGCTGTAAGGCGATGACTTCAAGAGAGAGCGCAACCTGCGGCGAC 420
Qy 421 AAGCTGAGTCAACCTAACAAGCAAGCCGCGCGAGGTGAAGTTCAGAGGCGAACAAGAAC 480
Db 421 AAGCTGAGTCAACCTAACAAGCAAGCCGCGCGAGGTGAAGTTCAGAGGCGAACAAGAAC 480
Qy 481 GGCATCAAGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCTGCAAGTCCGC 540
Db 481 GGCATCAAGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCTGCAAGTCCGC 540
Qy 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Db 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Qy 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCGCAAGAGCGAGTCAATGCTC 660
Db 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCGCAAGAGCGAGTCAATGCTC 660
Qy 661 CTGCTGAGTTCGTAGACCGCGCGGATCACTTCGCGATGAGCAAGTGAACAAG 720
Db 661 CTGCTGAGTTCGTAGACCGCGCGGATCACTTCGCGATGAGCAAGTGAACAAG 720
Qy 721 CTGAGCATGCTTCCTCCCGCGAGGTGAGAGAGAGAGATGAGACGCTGCCATGCT 780
Db 721 CTGAGCATGCTTCCTCCCGCGAGGTGAGAGAGAGAGATGAGACGCTGCCATGCT 780
Qy 781 TGTGCGCAGAGAGAGGAGGATGAGACCGTCACTTCGAGCTGCTTCTAGATCAAT 840
Db 781 TGTGCGCAGAGAGAGGAGGATGAGACCGTCACTTCGAGCTGCTTCTAGATCAAT 840
Qy 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 3
US-09-770-315-1
; Sequence 1, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7015
; TYPE: DNA
; ORGANISM: Unknown

FILE REFERENCE: NHL-ND-21
CURRENT APPLICATION NUMBER: US/09/503, 799
CURRENT FILING DATE: 2000-02-14

GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
APPLICANT: Ilodis, Juan

APPLICANT: Machter, Rebekka M.
 APPLICANT: Remington, S. James
 TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 FILE REFERENCE: 07257/071001
 CURRENT APPLICATION NUMBER: US/09/172,063
 EARLIER FILING DATE: 1998-10-13
 EARLIER APPLICATION NUMBER: 09/094,359
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 720
 TYPE: DNA
 ORGANISM: Aequorea victoria
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (0)..(0)
 OTHER INFORMATION: EGFP
 US-09-172-063-11

Query Match 84.4%; Score 717.4; DB 3; Length 720;
 Best Local Similarity 99.9%; Pred. No. 2.7e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTGAGCAGAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGAGCTGAGC 60
 1 ATGTGAGCAGAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGAGCTGAGC 60
 61 GCGCAGCTAAACGGCCACAAAGTTCAAGCGGTGCGGAGAGGGGCGCATCCCACTTAC 120
 61 GCGCAGCTAAACGGCCACAAAGTTCAAGCGGTGCGGAGAGGGGCGCATCCCACTTAC 120
 121 GCGAAGCTGACCCCTGAAGTTCACTGACACACCGGACAGTCCCGTGGCTGACCAAC 180
 121 GCGAAGCTGACCCCTGAAGTTCACTGACACACCGGACAGTCCCGTGGCTGACCAAC 180
 181 CTGTGACCAACCTGACCTTCAAGTCCGAGTGTCAAGCCGCTACCCGACCAATGAG 240
 181 CTGTGACCAACCTGACCTTCAAGTCCGAGTGTCAAGCCGCTACCCGACCAATGAG 240
 241 CAGCAGCACTTCTTCAAGTCCGAGTGTCAAGCCGAGTGTCAAGCCGAGTGTCAAG 300
 241 CAGCAGCACTTCTTCAAGTCCGAGTGTCAAGCCGAGTGTCAAGCCGAGTGTCAAG 300
 301 TTCAAGCAGAGCGGCACTAACAAGACCGGCGGAGTGTCAAGCCGAGTGTCAAG 360
 301 TTCAAGCAGAGCGGCACTAACAAGACCGGCGGAGTGTCAAGCCGAGTGTCAAG 360
 361 GTGAACCGCATCGAGCTGAGAGGGGATCGACTTCAAGAGAGAGGCAATCTTGGGCA 420
 361 GTGAACCGCATCGAGCTGAGAGGGGATCGACTTCAAGAGAGAGGCAATCTTGGGCA 420
 421 AAGCTGAGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 480
 421 AAGCTGAGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 480
 481 GGCATCAAGGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 540
 481 GGCATCAAGGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 540
 541 GACCACTACCAAGCAACAACCCCATCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 600
 541 GACCACTACCAAGCAACAACCCCATCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 600
 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 661 CTGCTGAGAGTTGAGACCGCGCGCGGAGATCACTCTCGGAGTGAAGAGAGTGAAGAA 719
 661 CTGCTGAGAGTTGAGACCGCGCGCGGAGATCACTCTCGGAGTGAAGAGAGTGAAGAA 719

RESULT 6
 US-09-316-919-12
 Sequence 12, Application US/09316919
 Patent No. 6469154
 GENERAL INFORMATION:
 APPLICANT: Telen, Roger Y.
 APPLICANT: Baird, Geoffrey
 TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS
 FILE REFERENCE: 07257/073001
 CURRENT APPLICATION NUMBER: US/09/316,919
 NUMBER OF SEQ ID NOS: 63
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12
 LENGTH: 720
 TYPE: DNA
 ORGANISM: Aequorea victoria
 US-09-316-919-12

Query Match 84.4%; Score 717.4; DB 4; Length 720;
 Best Local Similarity 99.9%; Pred. No. 2.7e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTGAGCAGAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGAGCTGAGC 60
 1 ATGTGAGCAGAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGAGCTGAGC 60
 61 GCGCAGCTAAACGGCCACAAAGTTCAAGCGGTGCGGAGAGGGGCGCATCCCACTTAC 120
 61 GCGCAGCTAAACGGCCACAAAGTTCAAGCGGTGCGGAGAGGGGCGCATCCCACTTAC 120
 121 GCGAAGCTGACCCCTGAAGTTCACTGACACACCGGACAGTCCCGTGGCTGACCAAC 180
 121 GCGAAGCTGACCCCTGAAGTTCACTGACACACCGGACAGTCCCGTGGCTGACCAAC 180
 181 CTGTGACCAACCTGACCTTCAAGTCCGAGTGTCAAGCCGCTACCCGACCAATGAG 240
 181 CTGTGACCAACCTGACCTTCAAGTCCGAGTGTCAAGCCGCTACCCGACCAATGAG 240
 241 CAGCAGCACTTCTTCAAGTCCGAGTGTCAAGCCGAGTGTCAAGCCGAGTGTCAAG 300
 241 CAGCAGCACTTCTTCAAGTCCGAGTGTCAAGCCGAGTGTCAAGCCGAGTGTCAAG 300
 301 TTCAAGCAGAGCGGCACTAACAAGACCGGCGGAGTGTCAAGCCGAGTGTCAAG 360
 301 TTCAAGCAGAGCGGCACTAACAAGACCGGCGGAGTGTCAAGCCGAGTGTCAAG 360
 361 GTGAACCGCATCGAGCTGAGAGGGGATCGACTTCAAGAGAGAGGCAATCTTGGGCA 420
 361 GTGAACCGCATCGAGCTGAGAGGGGATCGACTTCAAGAGAGAGGCAATCTTGGGCA 420
 421 AAGCTGAGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 480
 421 AAGCTGAGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 480
 481 GGCATCAAGGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 540
 481 GGCATCAAGGTGAACCTTCAAGATCGGCAACAAGCTTATCATGAGCGGACAAAGAAAC 540
 541 GACCACTACCAAGCAACAACCCCATCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 600
 541 GACCACTACCAAGCAACAACCCCATCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 600
 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 661 CTGCTGAGAGTTGAGACCGCGCGCGGAGATCACTCTCGGAGTGAAGAGAGTGAAGAA 719
 661 CTGCTGAGAGTTGAGACCGCGCGCGGAGATCACTCTCGGAGTGAAGAGAGTGAAGAA 719

481 GGCACTCAGGTAACTTCAGATCCGGCACACACTTGAAGA CGGACCGTGTAAGCTGCCG 540
|||||
|||

Db 493 CAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGAGCGACCATCTTC 552
 QY 301 TTCAAGAGCAGACGGCACTACAAAGACCCCGCGCGAGTGAAGTTGAGGGGACACACCTTG 360
 Db 553 TTCAAGAGCAGACGGCACTACAAAGACCCCGCGCGAGTGAAGTTGAGGGGACACACCTTG 612
 QY 361 GTGAACCCGATCGAGCTGAAGGGCATTCGATTCAGAGAGAGCGACATCTCTGGGGGAC 420
 Db 613 GTGAACCCGATCGAGCTGAAGGGCATTCGATTCAGAGAGAGCGACATCTCTGGGGGAC 672
 QY 421 AAGCTGAGTCACTAACAAGCAGCAGCAACGCTCTATCATAGGCGGCAACAAGCAGAGAC 480
 Db 673 AAGCTGAGTCACTAACAAGCAGCAGCAACGCTCTATCATAGGCGGCAACAAGCAGAGAC 732
 QY 481 GGCATCAAGGTGAATTCAGATCCGCCAACAATCGAGAGCGAGCGTCAAGTCCGCC 540
 Db 733 GGCATCAAGGTGAATTCAGATCCGCCAACAATCGAGAGCGAGCGTCAAGTCCGCC 792
 QY 541 GACCACTACAGCAGCAAGACCCCATCGGCGAGCGCCCGTGTCTGCTCCCGACACAC 600
 Db 793 GACCACTACAGCAGCAAGACCCCATCGGCGAGCGCCCGTGTCTGCTCCCGACACAC 852
 QY 601 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGAGTTC 660
 Db 853 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGAGTTC 912
 QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATGATCTCTGCGCATGAGAGAGCTGTACAGAA 719
 Db 913 CTGCTGAGTTCGTGACCGCGCGCGGATGATCTCTGCGCATGAGAGAGCTGTACAGAA 971

RESULT 11
 US-09-085-305-5
 ; Sequence 5, Application US/09085305
 ; Patent No. 6191269
 ; GENERAL INFORMATION:

APPLICANT: Pollock, Allan
 APPLICANT: Lovett, David H.
 TITLE OF INVENTION: Selective Induction of Apoptosis in
 TITLE OF INVENTION: Malignant Cancer Cells by Delivery of N-Terminal
 TITLE OF INVENTION: Interleukin-1-Alpha Pro-Piece Polypeptide
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bozicevic & Reed, LLP
 STREET: 285 Hamilton Ave, Suite 200
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94301

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSHO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/085.305
 FILING DATE: 29-MAY-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Francis, Carol L.
 REGISTRATION NUMBER: 36,513
 REFERENCE/DOCKET NUMBER: 6510/102US1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-327-3400
 TELEFAX: 650-327-3231
 TELEX:
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1095 base pairs

TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 1...1092
 OTHER INFORMATION:
 US-09-085-305-5

Query Match 84.4%; Score 717.4; DB 3; Length 1095;
 Best Local Similarity 99.9%; Pred. No. 2.8e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGGGAGAGGCTGTGACCGGGGTGGTGGCCATCTCTGTGAGCTGAGAC 60
 Db 376 ATGGTGAAGCAAGGGGAGAGGCTGTGACCGGGGTGGTGGCCATCTCTGTGAGCTGAGAC 435
 QY 61 GGCAGCGTAAACGGGCAAAAGTTGAGGTGTCCGCGAGAGGCGAGAGGCGATGCGACCTAC 120
 Db 436 GGCAGCGTAAACGGGCAAAAGTTGAGGTGTCCGCGAGAGGCGAGAGGCGATGCGACCTAC 495
 QY 121 GGCAGCGTAAACGGGCAAAAGTTGAGGTGTCCGCGAGAGGCGAGAGGCGATGCGACCTAC 180
 Db 496 GGCAGCGTAAACGGGCAAAAGTTGAGGTGTCCGCGAGAGGCGAGAGGCGATGCGACCTAC 555
 QY 181 CTGCTGAGCAGCCCTGACCTGAGCGGCTGAGCTGAGCTGAGCGGCTGACCCCGACACATGAG 240
 Db 556 CTGCTGAGCAGCCCTGACCTGAGCGGCTGAGCTGAGCTGAGCGGCTGACCCCGACACATGAG 615
 QY 241 CAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGAGCGCACCATCTTC 300
 Db 616 CAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGAGCGCACCATCTTC 675
 QY 301 TTCAAGAGCAGACGGCACTACAAAGACCCCGCGAGTGAAGTTGAGGGGACACACCTTG 360
 Db 676 TTCAAGAGCAGACGGCACTACAAAGACCCCGCGAGTGAAGTTGAGGGGACACACCTTG 735
 QY 361 GTGAACCCGATCGAGCTGAAGGGGATGATCTTCAAGAGAGAGCGGCAACATCTGAGGGGAC 420
 Db 736 GTGAACCCGATCGAGCTGAAGGGGATGATCTTCAAGAGAGAGCGGCAACATCTGAGGGGAC 795
 QY 421 AAGCTGAGTCACTAACAAGCAGCAGCAACGCTCTATCATAGGCGGCAACAAGCAGAGAAC 480
 Db 796 AAGCTGAGTCACTAACAAGCAGCAGCAACGCTCTATCATAGGCGGCAACAAGCAGAGAAC 855
 QY 481 GGCATCAAGGTGAATTCAGATCCGCCAACAATCGAGAGCGGATGAGCTGAGCTGCC 540
 Db 856 GGCATCAAGGTGAATTCAGATCCGCCAACAATCGAGAGCGGATGAGCTGAGCTGCC 915
 QY 541 GACCACTACAGCAGCAAGACCCCATCGGCGAGCGCCCGTGTCTGCTCCCGACACAC 600
 Db 916 GACCACTACAGCAGCAAGACCCCATCGGCGAGCGCCCGTGTCTGCTCCCGACACAC 975
 QY 601 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGAGTTC 660
 Db 976 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGAGTTC 1035
 QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATGATCTCTGCGCATGAGAGAGCTGTACAGAA 719
 Db 1036 CTGCTGAGTTCGTGACCGCGCGCGGATGATCTCTGCGCATGAGAGAGCTGTACAGAA 1094

RESULT 12
 US-09-417-197-128
 ; Sequence 128, Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 APPLICANT: Ole THASTRUP, et al.
 TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
 TITLE OF INVENTION: On A Cellular Response
 FILE REFERENCE: 3759-0110P
 CURRENT APPLICATION NUMBER: US/09/417,197

;; CURRENT FILING DATE: 1999-10-07
;; NUMBER OF SEQ ID NOS: 143
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 128
;; LENGTH: 1140
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: actin-binding-domain-Egfp fusion
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(1137)
US-09-417-197-128

Query Match 84.4%; Score 717.4; DB 4; Length 1140;
Best Local Similarity 99.9%; Pred. No. 2.9e-134;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGCTGTTCAACGGGGTGTGCGCCATCTGTGAGCTGAGC 60
Db 421 ATGTGAGCAAGGGGAGAGCTGTTCAACGGGGTGTGCGCCATCTGTGAGCTGAGC 480
QY 61 GGCGACGTAAACGGCCACAAGTTCAAGCGTCCGGGAGGGGAGGGGAGTCCCACTTAC 120
Db 481 GGCGACGTAAACGGCCACAAGTTCAAGCGTCCGGGAGGGGAGGGGAGTCCCACTTAC 540
QY 121 GGCAAGCTGACCTTGAAGTTCTATCTGACACACCGGCAAGCTGCGCTGCGCCAC 180
Db 541 GGCAAGCTGACCTTGAAGTTCTATCTGACACACCGGCAAGCTGCGCTGCGCCAC 600
QY 181 CTCGAGACCACTTCACTGAGGGGTGAGTCTTCAAGCGCTTACCCGACCAATGAG 240
Db 601 CTCGAGACCACTTCACTGAGGGGTGAGTCTTCAAGCGCTTACCCGACCAATGAG 660
QY 241 CAGCAGCACTTCTTAAGTCCGGCCATGCGCCGAGGCTAAGTCCAGAGGCACTCTTC 300
Db 661 CAGCAGCACTTCTTAAGTCCGGCCATGCGCCGAGGCTAAGTCCAGAGGCACTCTTC 720
QY 301 TTCAAGAGCGACGGGCACTTCAAGACCGCGCGGAGGTTGAAGTTGAGGGGCACTCTG 360
Db 721 TTCAAGAGCGACGGGCACTTCAAGACCGCGCGGAGGTTGAAGTTGAGGGGCACTCTG 780
QY 361 GTGAACCGCATCGAGCTGAGGAGTCACTTCAAGAGGAGACGGCAATCTCTGGGAC 420
Db 781 GTGAACCGCATCGAGCTGAGGAGTCACTTCAAGAGGAGACGGCAATCTCTGGGAC 840
QY 421 AAGCTGAGTCAACTACAGACGCAACGCTTATATCATGCGCGACAGACGAGAAC 480
Db 841 AAGCTGAGTCAACTACAGACGCAACGCTTATATCATGCGCGACAGACGAGAAC 900
QY 481 GGCACTAAGGTGAATTTCAAGTCCGCAACAATCGAGACGGGAGGCTGAGCTGCC 540
Db 901 GGCACTAAGGTGAATTTCAAGTCCGCAACAATCGAGACGGGAGGCTGAGCTGCC 960
QY 541 GACCACTACAGAGCAACCCCATCGGCGACCGGCGCGGCTGCGCCGAGCAACAC 600
Db 961 GACCACTACAGAGCAACCCCATCGGCGACCGGCGCGGCTGCGCCGAGCAACAC 1020
QY 601 TACCTGAGCACTTCAAGTCCGCTTGAAGAAACCCCAAGAGGCGGATCAATGCTC 660
Db 1021 TACCTGAGCACTTCAAGTCCGCTTGAAGAAACCCCAAGAGGCGGATCAATGCTC 1080
QY 661 CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAAGTA 719
Db 1081 CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAAGTA 1139

RESULT 13
US-09-417-197-112
; Sequence 112 Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.

;; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
;; FILE REFERENCE: 3759-0110P
;; CURRENT APPLICATION NUMBER: US/09/417,197
;; CURRENT FILING DATE: 1999-10-07
;; NUMBER OF SEQ ID NOS: 143
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 112
;; LENGTH: 1635
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: CDK2-Egfp fusion
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(1632)
US-09-417-197-112

Query Match 84.4%; Score 717.4; DB 4; Length 1635;
Best Local Similarity 99.9%; Pred. No. 3e-134;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGCTGTTCAACGGGGTGTGCGCCATCTGTGAGCTGAGC 60
Db 916 ATGTGAGCAAGGGGAGAGCTGTTCAACGGGGTGTGCGCCATCTGTGAGCTGAGC 975
QY 61 GGCGACGTAAACGGCCACAAGTTCAAGCGTCCGGGAGGGGAGGGGAGTCCCACTTAC 120
Db 976 GGCGACGTAAACGGCCACAAGTTCAAGCGTCCGGGAGGGGAGGGGAGTCCCACTTAC 1035
QY 121 GGCAAGCTGACCTTGAAGTTCTATCTGACACACCGGCAAGCTGCGCTGCGCCAC 180
Db 1036 GGCAAGCTGACCTTGAAGTTCTATCTGACACACCGGCAAGCTGCGCTGCGCCAC 1095
QY 181 CTCGAGACCACTTCACTGAGGGGTGAGTCTTCAAGCGCTTACCCGACCAATGAG 240
Db 1096 CTCGAGACCACTTCACTGAGGGGTGAGTCTTCAAGCGCTTACCCGACCAATGAG 1155
QY 241 CAGCAGCACTTCTTAAGTCCGGCCATGCGCCGAGGCTAAGTCCAGAGGCACTCTTC 300
Db 1156 CAGCAGCACTTCTTAAGTCCGGCCATGCGCCGAGGCTAAGTCCAGAGGCACTCTTC 1215
QY 301 TTCAAGAGCGACGGGCACTTCAAGACCGCGCGGAGGTTGAAGTTGAGGGGCACTCTG 360
Db 1216 TTCAAGAGCGACGGGCACTTCAAGACCGCGCGGAGGTTGAAGTTGAGGGGCACTCTG 1275
QY 361 GTGAACCGCATCGAGCTGAGGAGTCACTTCAAGAGGAGACGGCAATCTCTGGGAC 420
Db 1276 GTGAACCGCATCGAGCTGAGGAGTCACTTCAAGAGGAGACGGCAATCTCTGGGAC 1335
QY 421 AAGCTGAGTCAACTACAGACGCAACGCTTATATCATGCGCGACAGACGAGAAC 480
Db 1336 AAGCTGAGTCAACTACAGACGCAACGCTTATATCATGCGCGACAGACGAGAAC 1395
QY 481 GGCACTAAGGTGAATTTCAAGTCCGCAACAATCGAGACGGGAGGCTGAGCTGCC 540
Db 1396 GGCACTAAGGTGAATTTCAAGTCCGCAACAATCGAGACGGGAGGCTGAGCTGCC 1455
QY 541 GACCACTACAGAGCAACCCCATCGGCGACCGGCGCGGCTGCGCCGAGCAACAC 600
Db 1456 GACCACTACAGAGCAACCCCATCGGCGACCGGCGCGGCTGCGCCGAGCAACAC 1515
QY 601 TACCTGAGCACTTCAAGTCCGCTTGAAGAAACCCCAAGAGGCGGATCAATGCTC 660
Db 1516 TACCTGAGCACTTCAAGTCCGCTTGAAGAAACCCCAAGAGGCGGATCAATGCTC 1575
QY 661 CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAAGTA 719
Db 1576 CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAAGTA 1634

RESULT 14
US-09-417-197-58

Sequence 58, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 58
LENGTH: 1815
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Etk2-BGFP fusion
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1812)
US-09-417-197-58

Query Match 84.4%; Score 717.4; DB 4; Length 1815;
Best Local Similarity 99.9%; Pred. No. 3e-134; Indels 0; Gaps 0;
Matches 718; Conservative 0; Mismatches 1;

QY 1 ATGCTGACGAAAGGCGAGAGCTGTTCAACCGGGGTGTGCCCCATCTGTGAGCTGAGC 60
Db 1096 ATGCTGACGAAAGGCGAGAGCTGTTCAACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1155
QY 61 GGCACGCTAAACGGCGCAAAAGTTGAGCGGTGTCGGCGAGCCTGAGGCGATGCCACTTAC 120
Db 1156 GGCACGCTAAACGGCGCAAAAGTTGAGCGGTGTCGGCGAGGCGATGCCACTTAC 1215
QY 121 GGCAGCTGACCTCTGAAGTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 180
Db 1216 GGCAGCTGACCTCTGAAGTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 1275
QY 181 CTCTGACCACTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 240
Db 1276 CTCTGACCACTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 1335
QY 241 CAGACGACCTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 300
Db 1336 CAGACGACCTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 1395
QY 301 TTCAAGGACGACGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 360
Db 1396 TTCAAGGACGACGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 1455
QY 361 GTGAACCGCATGAGCTGAGGCGCATGACTTCAAGGAGGACGCGCAACATCTGCGGCGAC 420
Db 1456 GTGAACCGCATGAGCTGAGGCGCATGACTTCAAGGAGGACGCGCAACATCTGCGGCGAC 1515
QY 421 AAGCTGAGTAACTCACTCAAGCAAGCAACGCTCTATTCATGCGGCAAGCAAGCAAGAC 480
Db 1516 AAGCTGAGTAACTCACTCAAGCAAGCAACGCTCTATTCATGCGGCAAGCAAGCAAGAC 1575
QY 481 GGCATCAAGTGAATCTCAAGATCCGCGCAACATCGAGGACGCGCAAGCTGAGCTGCGC 540
Db 1576 GGCATCAAGTGAATCTCAAGATCCGCGCAACATCGAGGACGCGCAAGCTGAGCTGCGC 1635
QY 541 GACCACTAACGAGCAACACCCCATGCGGCGAGCGCGCGCTGCTGCGCGACACAC 600
Db 1636 GACCACTAACGAGCAACACCCCATGCGGCGAGCGCGCGCTGCTGCGCGACACAC 1695
QY 601 TACCTGACACCCAGTCCGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 660
Db 1696 TACCTGACACCCAGTCCGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 1755
QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTGCGCATGAGCGAGCTGTACAGATA 719
Db 1756 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTGCGCATGAGCGAGCTGTACAGATA 1814

RESULT 15
US-09-417-197-64
Sequence 64, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 64
LENGTH: 1821
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: p38-BGFP fusion
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1818)
US-09-417-197-64

Query Match 84.4%; Score 717.4; DB 4; Length 1821;
Best Local Similarity 99.9%; Pred. No. 3e-134; Indels 0; Gaps 0;
Matches 718; Conservative 0; Mismatches 1;

QY 1 ATGCTGACGAAAGGCGAGAGCTGTTCAACCGGGGTGTGCCCCATCTGTGAGCTGAGC 60
Db 1102 ATGCTGACGAAAGGCGAGAGCTGTTCAACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1161
QY 61 GGCACGCTAAACGGCGCAAAAGTTGAGCGGTGTCGGCGAGCCTGAGGCGATGCCACTTAC 120
Db 1162 GGCACGCTAAACGGCGCAAAAGTTGAGCGGTGTCGGCGAGGCGATGCCACTTAC 1221
QY 121 GGCAGCTGACCTCTGAAGTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 180
Db 1222 GGCAGCTGACCTCTGAAGTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 1281
QY 181 CTCTGACCACTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 240
Db 1282 CTCTGACCACTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 1341
QY 241 CAGACGACCTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 300
Db 1342 CAGACGACCTTCTCAAGTCCGCGATGCGGTGAGTCTTCAAGCGCTACCCGCAACATGAG 1401
QY 301 TTCAAGGACGACGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 360
Db 1402 TTCAAGGACGACGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 1461
QY 361 GTGAACCGCATGAGCTGAGGCGCATGACTTCAAGGAGGACGCGCAACATCTGCGGCGAC 420
Db 1462 GTGAACCGCATGAGCTGAGGCGCATGACTTCAAGGAGGACGCGCAACATCTGCGGCGAC 1521
QY 421 AAGCTGAGTAACTCACTCAAGCAAGCAACGCTCTATTCATGCGGCAAGCAAGCAAGAC 480
Db 1522 AAGCTGAGTAACTCACTCAAGCAAGCAACGCTCTATTCATGCGGCAAGCAAGCAAGAC 1581
QY 481 GGCATCAAGTGAATCTCAAGATCCGCGCAACATCGAGGACGCGCAAGCTGAGCTGCGC 540
Db 1582 GGCATCAAGTGAATCTCAAGATCCGCGCAACATCGAGGACGCGCAAGCTGAGCTGCGC 1641
QY 541 GACCACTAACGAGCAACACCCCATGCGGCGAGCGCGCGCTGCTGCGCGACACAC 600
Db 1642 GACCACTAACGAGCAACACCCCATGCGGCGAGCGCGCGCTGCTGCGCGACACAC 1701
QY 601 TACCTGACACCCAGTCCGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 660
Db 1702 TACCTGACACCCAGTCCGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 1761

Mon May 17 09:20:38 2004

us-09-931-232-2.rtf

Page 10

QY	CTGCTGGAGTCTGTAACCGCCGCCGGATCACTCCGCACTGGACGAGCTGTACAAGAA	719
Db	CTGCTGGAGTCTGTAACCGCCGCCGGATCACTCCGCACTGGACGAGCTGTACAAGTA	1820

Search completed: May 15, 2004, 05:22:04
Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 04:33:32 ; Search time 465 Seconds

(without alignments)
8295.358 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850
1 atggtgagcaaggcgagga.....taggatcaatgtgagatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapept 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	850	100.0	850	US-09-931-232-2	Sequence 2, Appli
2	850	100.0	4862	US-10-161-403-87	Sequence 87, Appli
3	721.2	84.8	10417	US-10-152-040-28	Sequence 28, Appli
4	720.6	84.8	6418	US-10-408-456-2	Sequence 2, Appli
5	720	84.7	12789	US-10-666-778-9	Sequence 9, Appli
6	719.4	84.6	8528	US-10-239-804-11	Sequence 11, Appli
7	719.2	84.6	5713	US-09-963-206B-5	Sequence 5, Appli
8	719.2	84.6	5713	US-09-966-976A-5	Sequence 5, Appli
9	719	84.6	8531	US-10-429-608A-2	Sequence 2, Appli
10	719	84.6	8531	US-10-134-643-5	Sequence 5, Appli
11	719	84.6	8531	US-10-408-456-1	Sequence 1, Appli
12	719	84.6	8531	US-10-716-725-2	Sequence 2, Appli
13	718.6	84.5	1190	US-09-927-876-98	Sequence 98, Appli
14	718.6	84.5	1190	US-10-457-047-98	Sequence 98, Appli

15	718.6	84.5	1190	US-10-360-149-98	Sequence 98, Appli
16	718.6	84.5	1377	US-10-204-724-1	Sequence 1, Appli
17	718.6	84.5	4151	US-10-421-285-15	Sequence 15, Appli
18	718.6	84.5	4733	US-09-797-496B-1	Sequence 1, Appli
19	718.6	84.5	4733	US-10-177-590-1	Sequence 1, Appli
20	718.6	84.5	4944	US-10-001-189-55	Sequence 55, Appli
21	718.6	84.5	4952	US-10-001-189-52	Sequence 52, Appli
22	718.6	84.5	9041	US-10-001-189-50	Sequence 50, Appli
23	718.6	84.5	9041	US-10-243-553-3	Sequence 3, Appli
24	718.6	84.5	9941	US-10-243-816-2	Sequence 2, Appli
25	718.6	84.5	9941	US-10-243-820-9	Sequence 9, Appli
26	718.6	84.5	9941	US-10-243-817-9	Sequence 9, Appli
27	718.6	84.5	9941	US-10-319-341-3	Sequence 3, Appli
28	718.2	84.5	5162	US-09-815-981-13	Sequence 13, Appli
29	718.2	84.5	5162	US-09-815-979-13	Sequence 13, Appli
30	718.2	84.5	5162	US-10-428-653-13	Sequence 13, Appli
31	718.2	84.5	5162	US-10-325-119-13	Sequence 13, Appli
32	718.2	84.5	5162	US-10-161-403-26	Sequence 26, Appli
33	718.2	84.5	5162	US-10-086-745-13	Sequence 13, Appli
34	718	84.5	5100	US-09-954-483A-14	Sequence 14, Appli
35	718	84.5	5510	US-10-161-403-71	Sequence 71, Appli
36	718	84.5	6100	US-10-314-861-36	Sequence 36, Appli
37	718	84.5	6115	US-10-314-861-34	Sequence 34, Appli
38	718	84.5	6119	US-10-161-403-126	Sequence 126, Appli
39	718	84.5	6121	US-10-314-861-38	Sequence 38, Appli
40	718	84.5	6133	US-10-314-861-32	Sequence 32, Appli
41	718	84.5	6151	US-10-314-861-30	Sequence 30, Appli
42	718	84.5	6169	US-10-314-861-28	Sequence 28, Appli
43	718	84.5	6248	US-10-314-861-15	Sequence 15, Appli
44	718	84.5	7600	US-10-161-403-115	Sequence 115, Appli
45	718	84.5	7631	US-10-161-403-116	Sequence 116, Appli

ALIGNMENTS

RESULT 1
US-09-931-232-2
; Sequence 2, Application US/09931232
; Publication No. US20020058274A1
GENERAL INFORMATION:
APPLICANT: Li, Xiangqiang
APPLICANT: Kahn, Steve
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
TITLE OF INVENTION: and Methods of Use
FILE REFERENCE: CLON075CON
CURRENT APPLICATION NUMBER: US/09/931,232
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/364,946
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/191,233
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 09/062,102
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 60/060,855
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of the EGFP-MOD422_461 fusion
OTHER INFORMATION: protein.

US-09-931-232-2

Query Match 100.0%; Score 850; DB 13; Length 850;
Best Local Similarity 100.0%; Pred. No. 6,4e-190;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGAGCAAGGCGAGGAGCTGTTCACCGGGGTGTGTCCTGCTGAGCTGAC 60
|||||


```

Db      1 ATGGTGAAGCAAGGCGAGAGAGCTGTTACCGGGGTGTGTGCCATCTGTGTGAGCTGAGC 60
QY      61 GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 120
Db      61 GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 120
QY      121 GGGAGAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 180
Db      121 GGGAGAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 180
QY      181 CTGTGTGAACCACTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 240
Db      181 CTGTGTGAACCACTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 240
QY      241 CAGCAGCACTTCTTCAAGTCCGCGATGCGCGAAGGCTGCAAGTCCAGAGGCGCACTTTC 300
Db      241 CAGCAGCACTTCTTCAAGTCCGCGATGCGCGAAGGCTGCAAGTCCAGAGGCGCACTTTC 300
QY      301 TTCAAGGACGACGCGCAACTCAAAAGACCGCGCGGAGGTGAAGTTTGAAGGCGACACCTTG 360
Db      301 TTCAAGGACGACGCGCAACTCAAAAGACCGCGCGGAGGTGAAGTTTGAAGGCGACACCTTG 360
QY      361 GTGAACCGCATTCAGCTGAAGGGGATGCACTTCAAGAGAGACGGCAACATCTTGGGGCAC 420
Db      361 GTGAACCGCATTCAGCTGAAGGGGATGCACTTCAAGAGAGACGGCAACATCTTGGGGCAC 420
QY      421 AAGCTGAGTAACTTCAACAGCAAGCAAGCTCTATCTATCTATCTATCTATCTATCTATCT 480
Db      421 AAGCTGAGTAACTTCAACAGCAAGCAAGCTCTATCTATCTATCTATCTATCTATCTATCT 480
QY      481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGGACGCGACGCTGACGCTCGCC 540
Db      481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGGACGCGACGCTGACGCTCGCC 540
QY      541 GACCACTTACAGAGAAACACCCCATTCGGGAGCGGCGCGTGTCTGTCTGTCTGTCTGTCT 600
Db      541 GACCACTTACAGAGAAACACCCCATTCGGGAGCGGCGCGTGTCTGTCTGTCTGTCTGTCT 600
QY      601 TACCTGAGCAACCGGATCCGCGTGAAGAAAGACCCCAACGAGAGCGGATCACTGTGTC 660
Db      601 TACCTGAGCAACCGGATCCGCGTGAAGAAAGACCCCAACGAGAGCGGATCACTGTGTC 660
QY      661 CTGCTGAGTGTCTGTAACCGCGCGCGGATCACTGTGAGCAAGCTGTCAAGAAAG 720
Db      661 CTGCTGAGTGTCTGTAACCGCGCGCGGATCACTGTGAGCAAGCTGTCAAGAAAG 720
QY      721 CTTAGCATGAGCTTCCCGCGGAGTGAAGAGCAAGATGAGCAAGCTGTGAGCAAGCTGT 780
Db      721 CTTAGCATGAGCTTCCCGCGGAGTGAAGAGCAAGATGAGCAAGCTGTGAGCAAGCTGT 780
QY      781 TGTGCCAGAGAGAGCGGATGAGACCGTCAACCTGCAAGCTGTCTGTCTGTGAGTCAAT 840
Db      781 TGTGCCAGAGAGAGCGGATGAGACCGTCAACCTGCAAGCTGTCTGTCTGTGAGTCAAT 840
QY      841 GTGTAGATGC 850
Db      841 GTGTAGATGC 850

```

RESULT 2

```

US-10-161-403-87
; Sequence 87, Application US/10161403
; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perez, Edward
; APPLICANT: Perkins, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS

```

```

; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10/161,403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 4862
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pDzeGFP-IN plasmid from Clontech
US-10-161-403-87

```

```

Query Match      100.0%; Score 850; DB 15; Length 4862;
Best Local Similarity 100.0%; Pred. No. 7.5e-150;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGGTGAAGCAAGGCGAGAGAGCTGTTACCGGGGTGTGTGCCATCTGTGTGAGCTGAGC 60
Db      679 ATGGTGAAGCAAGGCGAGAGAGCTGTTACCGGGGTGTGTGCCATCTGTGTGAGCTGAGC 738
QY      61 GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 120
Db      739 GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 798
QY      121 GGGAGAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 180
Db      799 GGGAGAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 858
QY      181 CTGTGTGAACCACTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 240
Db      859 CTGTGTGAACCACTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGGCCACAC 918
QY      241 CAGCAGCACTTCTTCAAGTCCGCGATGCGCGAAGGCTGCAAGTCCAGAGGCGCACTTTC 300
Db      919 CAGCAGCACTTCTTCAAGTCCGCGATGCGCGAAGGCTGCAAGTCCAGAGGCGCACTTTC 978
QY      301 TTCAAGGACGACGCGCAACTCAAAAGACCGCGCGGAGGTGAAGTTTGAAGGCGACACCTTG 360
Db      979 TTCAAGGACGACGCGCAACTCAAAAGACCGCGCGGAGGTGAAGTTTGAAGGCGACACCTTG 1038
QY      361 GTGAACCGCATTCAGCTGAAGGGGATGCACTTCAAGAGAGAGGCAACATCTGAGGGCAC 420
Db      1039 GTGAACCGCATTCAGCTGAAGGGGATGCACTTCAAGAGAGAGGCAACATCTGAGGGCAC 1098
QY      421 AAGCTGAGTAACTTCAACAGCAAGCAAGCTCTATCTATCTATCTATCTATCTATCTATCT 480
Db      1099 AAGCTGAGTAACTTCAACAGCAAGCAAGCTCTATCTATCTATCTATCTATCTATCTATCT 1158
QY      481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGGACGCGACGCTGACGCTCGCC 540
Db      1159 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGGACGCGACGCTGACGCTCGCC 1218
QY      541 GACCACTTACAGAGAAACACCCCATTCGGGAGCGGCGCGTGTCTGTCTGTGAGTCAAT 600
Db      1219 GACCACTTACAGAGAAACACCCCATTCGGGAGCGGCGCGTGTCTGTCTGTGAGTCAAT 1278
QY      601 TACCTGAGCAACCGGATCCGCGTGAAGAAAGACCCCAACGAGAGCGGATCACTGTGTC 660
Db      1279 TACCTGAGCAACCGGATCCGCGTGAAGAAAGACCCCAACGAGAGCGGATCACTGTGTC 1338
QY      661 CTGCTGAGTGTCTGTAACCGCGCGCGGATCACTGTGAGCAAGCTGTGAGCAAGCTGT 720
Db      1339 CTGCTGAGTGTCTGTAACCGCGCGCGGATCACTGTGAGCAAGCTGTGAGCAAGCTGT 1398
QY      721 CTTAGCATGAGCTTCCCGCGGAGTGAAGAGCAAGATGAGCAAGCTGTGAGCAAGCTGT 780
Db      1399 CTTAGCATGAGCTTCCCGCGGAGTGAAGAGCAAGATGAGCAAGCTGTGAGCAAGCTGT 1458

```

QY 781 TGTGCCAGAGAGGGGATGACCTGACCTGAGCTGTGCTTGTGATGATCAAT 840
DB 1459 TGTGCCAGAGAGGGGATGACCTGACCTGAGCTGTGCTTGTGATGATCAAT 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 3

US-10-152-040-28
Sequence 28, Application US/10152040
Publication No. US20030077251A1
GENERAL INFORMATION:
APPLICANT: ESCRIOU, NICOLAS
APPLICANT: VAN DER WERF, SYLVIE
APPLICANT: VIGNUZZI, MARCO
APPLICANT: GERBAUD, SYLVIE
TITLE OF INVENTION: REPLICONS DERIVED FROM POSITIVE STRAND RNA VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 03495, 0229-0000
CURRENT APPLICATION NUMBER: US/10/152,040
CURRENT FILING DATE: 2002-06-27
PRIORITY FILING DATE: 2001-05-23
PRIORITY FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 10417
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-152-040-28

Query Match 84.8%; Score 721.2; DB 15; Length 10417;
Best Local Similarity 97.6%; Pred. No. 1.3e-159;
Matches 732; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGCGAGAGCTGTCAACCGGGGTGATGCCCATCTGTGTGAGCTGAC 60
DB 743 ATGTGAGCAAGGGCGAGAGCTGTCAACCGGGGTGATGCCCATCTGTGTGAGCTGAC 802
QY 61 GGGCAGCTAAACGGCCACCAAGTTCAAGCTGTCCGGCAGAGGGCGAGGCGATGCCACTAC 120
DB 803 GGGCAGCTAAACGGCCACCAAGTTCAAGCTGTCCGGCAGAGGGCGAGGCGATGCCACTAC 862
QY 121 GGGCAGCTGACCTGTAAGTTCTGCAACAACCGGCAAGCTGCGCCGTGCGCCACC 180
DB 863 GGGCAGCTGACCTGTAAGTTCTGCAACAACCGGCAAGCTGCGCCGTGCGCCACC 922
QY 181 CTGTGAGCAACCCCTGACCTGACGGGTGACAGTCTTCAACCGGCTACCCCGACCAATTAAG 240
DB 923 CTGTGAGCAACCCCTGACCTGACGGGTGACAGTCTTCAACCGGCTACCCCGACCAATTAAG 982
QY 241 CAGCAGCACTTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGGCCCAACCATCTTC 300
DB 983 CAGCAGCACTTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGGCCCAACCATCTTC 1042
QY 301 TTCAAGAGCAACGGCAACTCAAGAACCCGCGCGAGGCTGAAATTGAGGGCGACACCTTG 360
DB 1043 TTCAAGAGCAACGGCAACTCAAGAACCCGCGCGAGGCTGAAATTGAGGGCGACACCTTG 1102
QY 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGCAACGGCAACATCTGGGGCAC 420
DB 1103 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGCAACGGCAACATCTGGGGCAC 1162
QY 421 AAGCTGAGTAACTAACAACAGCAACAGCTCTATATCAATGAGCGCAAGCAGAGAGAAC 480
DB 1163 AAGCTGAGTAACTAACAACAGCAACAGCTCTATATCAATGAGCGCAAGCAGAGAGAAC 1222

QY 481 GGCATCAAGTGAATCTTCAAGATCCGCCACACATCCAGAGAGCGGACGCTGACCTGCC 540
DB 1223 GGCATCAAGTGAATCTTCAAGATCCGCCACACATCCAGAGAGCGGACGCTGACCTGCC 1282
QY 541 GACCACTACCAAGCAACACCCCATGCGGACGAGGCGCCCTGCTGTGCGCCGACACCTAC 600
DB 1283 GACCACTACCAAGCAACACCCCATGCGGACGAGGCGCCCTGCTGTGCGCCGACACCTAC 1342
QY 601 TACCTGAGCAACCCGATCCGCGCTGAGCAAGACCCCAACAGAGAGCGGATCATAGTGC 660
DB 1343 TACCTGAGCAACCCGATCCGCGCTGAGCAAGACCCCAACAGAGAGCGGATCATAGTGC 1402
QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTCGGATGAGACGAGTGTACAAAG 720
DB 1403 CTGCTGAGTTCGTGACCGCGCGCGGATCACTCTCGGATGAGACGAGTGTACAAAG 1462
QY 721 CTGAGCATGAGCTTCCGCGGAGGTGAG 750
DB 1463 CTGAGCATGAGCTTCCGCGGAGGTGAG 1492

RESULT 4

US-10-408-456-2
Sequence 2, Application US/10408456
Publication No. US20040013648A1
GENERAL INFORMATION:
APPLICANT: KINGSMAN, et al., Alan John
TITLE OF INVENTION: Vector System
FILE REFERENCE: 674523-2016
CURRENT APPLICATION NUMBER: US/10/408,456
CURRENT FILING DATE: 2003-04-08
PRIORITY FILING DATE: PCT/GB01/04433
PRIORITY FILING DATE: 2001-10-15
PRIORITY FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 2
LENGTH: 6418
TYPE: DNA
ORGANISM: Equine Infectious Anemia Virus
US-10-408-456-2

Query Match 84.8%; Score 720.6; DB 16; Length 6418;
Best Local Similarity 96.8%; Pred. No. 1.7e-159;
Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGCGAGAGCTGTCAACCGGGGTGATGCCCATCTGTGTGAGCTGAC 60
DB 2690 ATGTGAGCAAGGGCGAGAGCTGTCAACCGGGGTGATGCCCATCTGTGTGAGCTGAC 2749
QY 61 GGGCAGCTAAACGGCCACCAAGTTCAAGCTGTCCGGCAGAGGGCGAGGCGATGCCACTAC 120
DB 2750 GGGCAGCTAAACGGCCACCAAGTTCAAGCTGTCCGGCAGAGGGCGAGGCGATGCCACTAC 2809
QY 121 GGCAGCTGACCTGTAAGTTCTGCAACAACCGGCAAGCTGCGCCGTGCGCCACC 180
DB 2810 GGCAGCTGACCTGTAAGTTCTGCAACAACCGGCAAGCTGCGCCGTGCGCCACC 2869
QY 181 CTGTGAGCAACCCCTGACCTGACGGGTGACAGTCTTCAACCGGCTACCCCGACCAATTAAG 240
DB 2870 CTGTGAGCAACCCCTGACCTGACGGGTGACAGTCTTCAACCGGCTACCCCGACCAATTAAG 2929
QY 241 CAGCAGCACTTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGGCCCAACCATCTTC 300
DB 2930 CAGCAGCACTTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGGCCCAACCATCTTC 2989
QY 301 TTCAAGAGCAACGGCAACTCAAGAACCCGCGCGAGGCTGAAATTGAGGGCGACACCTTG 360
DB 2990 TTCAAGAGCAACGGCAACTCAAGAACCCGCGCGAGGCTGAAATTGAGGGCGACACCTTG 3049
QY 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGCAACGGCAACATCTGGGGCAC 420

DB 3050 GTGAACCGATCGAGCTGAAGGGGCGATCGCTTCAAGAGAGCGCAACATCTCGGGGCGAC 3109
QY 421 AAGCTGAGTACAACTACAAAGCCAAAGCTTATATCATGGCCGAGCAAGAGAAAG 480
DB 3110 AAGCTGAGTACAACTACAAAGCCAAAGCTTATATCATGGCCGAGCAAGAGAAAG 3169
QY 481 GGCATCAAGGTGAATCTTCAAGATCGCCCAACATCGAGAGCGGAGCGTGCAGTGC 540
DB 3170 GGCATCAAGGTGAATCTTCAAGATCGCCCAACATCGAGAGCGGAGCGTGCAGTGC 3229
QY 541 GACCACTACAGAGAAACCCCACTCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 600
DB 3230 GACCACTACAGAGAAACCCCACTCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 3289
QY 601 TACCTGAGACCCAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 660
DB 3290 TACCTGAGACCCAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 3349
QY 661 CTGCTGAGGTTCTGTCAGCGCGCGCGGATCATCTCGGATGAGCGAGCTGTACAAAG 720
DB 3350 CTGCTGAGGTTCTGTCAGCGCGCGCGGATCATCTCGGATGAGCGAGCTGTACAAAG 3409
QY 721 CTGAGCATGGCTCCCGCGGAGGAGGAGAGAGAT 759
DB 3410 AGCGGCGGAGCTGTAGAGTCACTGCAAGAAATTCGAT 3448

RESULT 5

US-10-666-778-9/C
; Sequence 9, Application US/10666778
; Publication No. US20040064849A1
; GENERAL INFORMATION:
; APPLICANT: Goossens, Alain
; APPLICANT: Inze, Dirk
; TITLE OF INVENTION: THE USE OF GENES ENCODING MEMBRANE TRANSPORT PUMPS TO STIMULATE
; FILE REFERENCE: DI/ABC/V082
; CURRENT APPLICATION NUMBER: US/10/666,778
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: EP01201407.2
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 12789
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4772)..(4772)
; OTHER INFORMATION: n can be any base
US-10-666-778-9

Query Match 84.7%; Score 720; DB 13; Length 12789;
Best Local Similarity 100.0%; Pred. No. 2.5e-159;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGAGCAAGGGGAGAGAGCTTACCGGGGTGGTCCCATCTGTGTGACGCTGAC 60
DB 5797 ATGGTGAGCAAGGGGAGAGAGCTTACCGGGGTGGTCCCATCTGTGTGACGCTGAC 5738
QY 61 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCGCGAGAGGCGAGCGATGCCACTTAC 120
DB 5737 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCGCGAGAGGCGAGCGATGCCACTTAC 5678
QY 121 GGGAGCTACCTGTAAAGTTCACTGTGACCAACCGGAGAGCTGCGCTGAGCCACC 180
DB 5677 GGGAGCTACCTGTAAAGTTCACTGTGACCAACCGGAGAGCTGCGCTGAGCCACC 5618
QY 181 CTGTGACCAACCTGACTTACGGCGTGCAGTGTCTTACGCGGCTAACCCGACCACTGAAG 240

DB 5617 CTGTGACCAACCTGACTTACGGCGTGCAGTGTCTTACCGCTTACCCGACCACTGAAG 5558
QY 241 CAGCAGACTTCTTAAAGTCCGCATGCGCCGAAAGGCTGCTCGAGAGCGGACCATCTTC 300
DB 5557 CAGCAGACTTCTTAAAGTCCGCATGCGCCGAAAGGCTGCTCGAGAGCGGACCATCTTC 5498
QY 301 TTCAAGAGCAGACGGCACTTACAAAGACCGCGCCGAGGTGAATTGAGGGGACACCTTG 360
DB 5497 TTCAAGAGCAGACGGCACTTACAAAGACCGCGCCGAGGTGAATTGAGGGGACACCTTG 5438
QY 361 GTGAACCGGATGAGCTGTAAGGGGCTTGAATTCAAGAGAGAGCGCAACATCTGGGGGAC 420
DB 5437 GTGAACCGGATGAGCTGTAAGGGGCTTGAATTCAAGAGAGAGCGCAACATCTGGGGGAC 5378
QY 421 AAGCTGAGTACAACTTACAAAGAGCAACGCTTATCATGGCCGAGCAAGAGAAAG 480
DB 5377 AAGCTGAGTACAACTTACAAAGAGCAACGCTTATCATGGCCGAGCAAGAGAAAG 5318
QY 481 GGCATCAAGGTGAATCTTCAAGATCGCCCAACATCGAGAGCGGAGCGTGCAGCTGC 540
DB 5317 GGCATCAAGGTGAATCTTCAAGATCGCCCAACATCGAGAGCGGAGCGTGCAGCTGC 5258
QY 541 GACCACTACAGAGAAACCCCACTCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 600
DB 5257 GACCACTACAGAGAAACCCCACTCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 5198
QY 601 TACCTGAGACCCAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 660
DB 5197 TACCTGAGACCCAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 5138
QY 661 CTGCTGAGGTTCTGTCAGCGCGCGGATCATCTCGGATGAGCGAGCTGTACAAAG 720
DB 5137 CTGCTGAGGTTCTGTCAGCGCGCGGATCATCTCGGATGAGCGAGCTGTACAAAG 5078

RESULT 6

US-10-239-804-11
; Sequence 11, Application US/10239804
; Publication No. US20030053991A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; APPLICANT: Kingsman, Alan J
; APPLICANT: Maden, Malcolm
; APPLICANT: Corcoran, Jonathan PT
; TITLE OF INVENTION: Factor
; FILE REFERENCE: P009156MOCFH
; CURRENT APPLICATION NUMBER: US/10/239,804
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/GB00/01211
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 0024300.6
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 8528
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pONY86 5' cPPT
; OTHER INFORMATION: POS delTTS E1AV vector genome plasmid
US-10-239-804-11

Query Match 84.6%; Score 719.4; DB 15; Length 8528;
Best Local Similarity 97.2%; Pred. No. 3.4e-159;
Matches 732; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGTGGTCCCATCTGTGTGACGCTGAC 60
DB 2868 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGTGGTCCCATCTGTGTGACGCTGAC 2927
QY 61 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCGCGAGAGGCGAGCGATGCCACTTAC 120

Db 2928 GCGCGCTGAAACGGCCCAAGATTGACGCTGTCCGCGGAGGGGAGGCGGATGCCACTAC 2987
Qy 121 GCGAAGCTGACCCCTGAATTCATCTGCAACACCGGCAAGCTGCGCCCTGCGCCAC 180
Db 2988 GCGAAGCTGACCCCTGAATTCATCTGCAACACCGGCAAGCTGCGCCCTGCGCCAC 3047
Qy 181 CTCGTGACCAACCCCTGACCTACCGCGTGAAGTCTTCAGCCGCTACCCCGACCACTAAG 240
Db 3048 CTCGTGACCAACCCCTGACCTACCGCGTGAAGTCTTCAGCCGCTACCCCGACCACTAAG 3107
Qy 241 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTACGTCCAGAGCGACCATCTTC 300
Db 3108 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTACGTCCAGAGCGACCATCTTC 3167
Qy 301 TTCAAGGACGAGCGGCACTCAAGACCGCGCGGAGGTAAGTTGAGGCGGACACCTTG 360
Db 3168 TTCAAGGACGAGCGGCACTCAAGACCGCGCGGAGGTAAGTTGAGGCGGACACCTTG 3227
Qy 361 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGGAGCGGCAACATCTGCGGAC 420
Db 3228 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGGAGCGGCAACATCTGCGGAC 3287
Qy 421 AAGCTGAGTCACTCAACAGCCCAACGCTTATATCATGCGGACCAAGCAAGAAC 480
Db 3288 AAGCTGAGTCACTCAACAGCCCAACGCTTATATCATGCGGACCAAGCAAGAAC 3347
Qy 481 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGGAGGAGCGGAGGTGAGTCCGC 540
Db 3348 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGGAGGAGCGGAGGTGAGTCCGC 3407
Qy 541 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 600
Db 3408 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 3467
Qy 601 TACCTGAGACCCAGTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 660
Db 3468 TACCTGAGACCCAGTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 3527
Qy 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTACAAGAG 720
Db 3528 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTACAAGAG 3587
Qy 721 CTGAGCATGCTTCCGCGGAGGTGAGAGAG 753
Db 3588 AGCGCGCGGACTTGAAGTGAAGTCTGAGAGG 3620

RESULT 7
US-09-963-206B-5
; Sequence 5, Application US/09963206B
; Patent No. US20020123076A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan E.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE REFERENCE: A-66038-3/RMS/JTD/DLR
; CURRENT APPLICATION NUMBER: US/09/963,206B
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/076,624
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5
; LENGTH: 5713
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-963-206B-5

Query Match 84.6%; Score 719.2; DB 9; Length 5713;
Best Local Similarity 98.9%; Pred. No. 3,66-159;
Matches 724; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGGCGAGAGCTTTCACCGGGGTGTGCTTCATCTGTGTGAGCTGAC 60
Db 2150 ATGTGAGCAAGGGCGAGAGCTTTCACCGGGGTGTGCTTCATCTGTGTGAGCTGAC 2209
Qy 61 GCGCAAGTAAACCGGCAAGATTGAGCGTGCAGGCGGAGGCGGATGCCACTAC 120
Db 2210 GCGCAAGTAAACCGGCAAGATTGAGCGTGCAGGCGGAGGCGGATGCCACTAC 2269
Qy 121 GCGAAGCTGACCCCTGAATTCATCTGCAACACCGGCAAGCTGCCGCTGCGCCAC 180
Db 2270 GCGAAGCTGACCCCTGAATTCATCTGCAACACCGGCAAGCTGCCGCTGCGCCAC 2329
Qy 181 CTCGTGACCAACCCCTGACCTACCGGCTGAGTGTCTTCAAGCGCTACCCGACCAATGAG 240
Db 2330 CTCGTGACCAACCCCTGACCTACCGGCTGAGTGTCTTCAAGCGCTACCCGACCAATGAG 2389
Qy 241 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTACGTCCAGAGCGGACCATCTTC 300
Db 2390 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTACGTCCAGAGCGGACCATCTTC 2449
Qy 301 TTCAAGGACGAGCGGCACTCAAGACCGCGCGGAGGTAAGTTGAGGCGGACACCTTG 360
Db 2450 TTCAAGGACGAGCGGCACTCAAGACCGCGCGGAGGTAAGTTGAGGCGGACACCTTG 2509
Qy 361 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGGAGCGGCAACATCTGCGGAC 420
Db 2510 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGGAGCGGCAACATCTGCGGAC 2569
Qy 421 AAGCTGAGTCACTCAACAGCCCAACGCTTATATCATGCGGCAACAGCAAGAAC 480
Db 2570 AAGCTGAGTCACTCAACAGCCCAACGCTTATATCATGCGGCAACAGCAAGAAC 2629
Qy 481 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGGAGGAGCGGATCACTGCTC 540
Db 2630 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGGAGGAGCGGATCACTGCTC 2689
Qy 541 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 600
Db 2690 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 2749
Qy 601 TACCTGAGACCCAGTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 660
Db 2750 TACCTGAGACCCAGTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 2809
Qy 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTACAAGAG 720
Db 2810 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTACAAGAG 2869
Qy 721 CTGAGCATGCTTCCGCGGAGGTGAGAGAG 752
Db 2870 CTGAGCATGCTTCCGCGGAGGTGAGAGAG 2881

RESULT 8
US-09-966-976A-5
; Sequence 5, Application US/09966976A
; Patent No. US2002016849A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan E.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE REFERENCE: A-66038-4/RMS/JTD/DLR
; CURRENT APPLICATION NUMBER: US/09/966,976A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 09/076,624

PRIOR FILING DATE: 1998-05-12
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 5713
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-966-976A-5

Query Match 84.6%; Score 719.2; DB 9; Length 5713;
Best Local Similarity 98.9%; Pred. No. 3.0e-159;
Matches 724; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTCTGTCAGCTGAGAC 60
DB 2150 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTCTGTCAGCTGAGAC 2209
QY 61 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 120
DB 2210 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2269
QY 121 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 180
DB 2270 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2329
QY 181 CTGCTGAAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 240
DB 2330 CTGCTGAAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2389
QY 241 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGGCGAGGCGAGTCCACTTAC 300
DB 2390 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGGCGAGGCGAGTCCACTTAC 2449
QY 301 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 360
DB 2450 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2509
QY 361 GTGAACCGGATGAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 420
DB 2510 GTGAACCGGATGAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2569
QY 421 AAGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 480
DB 2570 AAGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2629
QY 481 GGCATCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 540
DB 2630 GGCATCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2689
QY 541 GACCACTACCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 600
DB 2690 GACCACTACCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2749
QY 601 TACCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 660
DB 2750 TACCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2809
QY 661 CTGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 720
DB 2810 CTGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2869
QY 721 CTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 780
DB 2870 CTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2881

RESULT 9
US-10-429-608A-2
Sequence 2, Application US/10429608A
Publication No. US20040071675A1
GENERAL INFORMATION:

APPLICANT: MAZARAKIS, NICHOLAS
APPLICANT: AZOUZ, MIMOUN
TITLE OF INVENTION: VECTOR SYSTEM
FILE REFERENCE: 674523-2017
CURRENT APPLICATION NUMBER: US/10/429, 608A
CURRENT FILING DATE: 2003-05-05
PRIOR APPLICATION NUMBER: PCT/GB01/04866
PRIOR FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: GB 0122238.9
PRIOR FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: GB 0102339.9
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: GB 0026943.1
PRIOR FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 8531
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: nucleotide construct PONTAG sequence
US-10-429-608A-2

Query Match 84.6%; Score 719; DB 12; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTCTGTCAGCTGAGAC 60
DB 2690 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTCTGTCAGCTGAGAC 2749
QY 61 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 120
DB 2750 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2809
QY 121 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 180
DB 2810 GCGGACGTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGAGTCCACTTAC 2869
QY 181 CTGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 240
DB 2870 CTGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 2929
QY 241 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGGCGAGGCGAGTCCACTTAC 300
DB 2930 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGGCGAGGCGAGTCCACTTAC 2989
QY 301 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 360
DB 2990 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3049
QY 361 GTGAACCGGATGAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 420
DB 3050 GTGAACCGGATGAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3109
QY 421 AAGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 480
DB 3110 AAGCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3169
QY 481 GGCATCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 540
DB 3170 GGCATCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3229
QY 541 GACCACTACCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 600
DB 3230 GACCACTACCAAGGTAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3289
QY 601 TACCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 660
DB 3290 TACCTGAAGTGAAGTGAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGTCCACTTAC 3349

QY 661 CTGCTGAGAGTTCTGTAACCGCCCGGAGATCATCTTCGGCATGAGACGAGCTGTACAAG 720
DB 3350 CTGCTGAGAGTTCTGTAACCGCCCGGAGATCATCTTCGGCATGAGACGAGCTGTACAAG 3409
QY 721 CTGAGCATGAGCTTCCCGCCGAGGTGAGG 751
DB 3410 AGCGCCGCGAGCTTCTAGAGTGCAGCTGCAGG 3440

RESULT 10

US-10-134-643-5
; Sequence 5, Application US/10134643
; Publication No. US20030113898A1
; GENERAL INFORMATION:
; APPLICANT: OLSEN, JOHN C.
; APPLICANT: MITROPHANOUS, KYRIACOS ANDREOU
; APPLICANT: ROHLI, JONATHAN
; APPLICANT: KINGSMAN, ALAN JOHN
; APPLICANT: ELLIARD, FIONA MARGARET
; TITLE OF INVENTION: METHODS FOR PRODUCING HIGH TITRE VECTORS AND
; FILE REFERENCE: 078883-0148
; CURRENT APPLICATION NUMBER: US/10/134,643
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/287,048
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleic acid pONY8g, vector genome plasmid
US-10-134-643-5

Query Match 84.6%; Score 719; DB 15; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCAGCTGAGC 60
DB 2690 ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCAGCTGAGC 2749
QY 61 GGCGACGTAAAGCGGCACAAAGTTTCAAGCGGTGTCGGCGAGGCGAGGCGATGCCACTAC 120
DB 2750 GGCGACGTAAAGCGGCACAAAGTTTCAAGCGGTGTCGGCGAGGCGAGGCGATGCCACTAC 2809
QY 121 GGCAAGCTGACCTTGAAGTTTATCTGCAACCGGCAAGCTGCGGTGCTGAGCCCAAC 180
DB 2810 GGCAAGCTGACCTTGAAGTTTATCTGCAACCGGCAAGCTGCGGTGCTGAGCCCAAC 2869
QY 181 CTGCGACCAACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 240
DB 2870 CTGCGACCAACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2929
QY 241 CAGACGACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 300
DB 2930 CAGACGACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2989
QY 301 TTCAAGGACGACGAGCACTAACAAGCCCGCGCGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
DB 2990 TTCAAGGACGACGAGCACTAACAAGCCCGCGCGAGGCTGAGGCTGAGGCTGAGGCTGAG 3049
QY 361 GTGAACCGCATCGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
DB 3050 GTGAACCGCATCGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 3109
QY 421 AAGCTGAGTAACTAACAAGCCCAAGCTTATATCATGCGGACCAAGGAGGAGAAC 480
DB 3110 AAGCTGAGTAACTAACAAGCCCAAGCTTATATCATGCGGACCAAGGAGGAGAAC 3169

QY 481 GGCATCAAGTGAATTCAAGATCCGCCCAACATCGAGGAGCGGAGCGTGAAGCTCGCC 540
DB 3170 GGCATCAAGTGAATTCAAGATCCGCCCAACATCGAGGAGCGGAGCGTGAAGCTCGCC 3229
QY 541 GACCACTACACAGCAAGCAACCCCAATCGCGAGCGCCCGCTGCTGCGCCCAACACAC 600
DB 3230 GACCACTACACAGCAAGCAACCCCAATCGCGAGCGCCCGCTGCTGCGCCCAACACAC 3289
QY 601 TACCTGAGCAACCAAGTCCGCTTGAAGCAAGCCCAACAGAGAGCGATCAATGTC 660
DB 3290 TACCTGAGCAACCAAGTCCGCTTGAAGCAAGCCCAACAGAGAGCGATCAATGTC 3349
QY 661 CTGCTGAGAGTTGTAACCGCGCGGAGTCACTCTCGAGATGAGAGAGTGTACAAGAG 720
DB 3350 CTGCTGAGAGTTGTAACCGCGCGGAGTCACTCTCGAGATGAGAGAGTGTACAAGAG 3409
QY 721 CTGAGCATGAGCTTCCCGCCGAGGTGAGG 751
DB 3410 AGCGCCGCGAGCTTCTAGAGTGCAGCTGCAGG 3440

RESULT 11

US-10-408-456-1
; Sequence 1, Application US/10408456
; Publication No. US20040013648A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; APPLICANT: KINGSMAN, et al., Alan John
; TITLE OF INVENTION: Vector System
; FILE REFERENCE: 674523-2016
; CURRENT APPLICATION NUMBER: US/10/408,456
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/04433
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 0024550.6
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 1
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Equine Infectious Anemia Virus
US-10-408-456-1

Query Match 84.6%; Score 719; DB 15; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCAGCTGAGC 60
DB 2690 ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCCATCTCTGTCAGCTGAGC 2749
QY 61 GGCGACGTAAAGCGGCACAAAGTTTCAAGCGGTGTCGGCGAGGCGAGGCGATGCCACTAC 120
DB 2750 GGCGACGTAAAGCGGCACAAAGTTTCAAGCGGTGTCGGCGAGGCGAGGCGATGCCACTAC 2809
QY 121 GGCAAGCTGACCTTGAAGTTTATCTGCAACCGGCAAGCTGCGGTGCTGAGCCCAAC 180
DB 2810 GGCAAGCTGACCTTGAAGTTTATCTGCAACCGGCAAGCTGCGGTGCTGAGCCCAAC 2869
QY 181 CTGCGACCAACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 240
DB 2870 CTGCGACCAACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2929
QY 241 CAGACGACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 300
DB 2930 CAGACGACCTTCAAGCTGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2989
QY 301 TTCAAGGACGACGAGCACTAACAAGCCCGCGCGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
DB 2990 TTCAAGGACGACGAGCACTAACAAGCCCGCGCGAGGCTGAGGCTGAGGCTGAGGCTGAG 3049
QY 361 GTGAACCGCATCGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420

Db 3050 GTAAACCGCATCGAGCTGAAGGGCATCGACTTCAAGAGAGACGGCAACATCTGGGGCAC 3109
QY 421 AAGCTGAGTACACTACACAGCAGCAACGTCATATCATGGCCGACAGAGAAAGAAC 480
Db 3110 AAGCTGAGTACACTACACAGCAGCAACGTCATATCATGGCCGACAGAGAAAGAAC 3169
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGCC 540
Db 3170 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGCC 3229
QY 541 GACCACTACACAGAGAACACCCCATCGGAGACGGCCGCTGCTGCTGCCGACAAACAC 600
Db 3230 GACCACTACACAGAGAACACCCCATCGGAGACGGCCGCTGCTGCTGCCGACAAACAC 3289
QY 601 TACCTGAGACCCAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 660
Db 3290 TACCTGAGACCCAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 3349
QY 661 CTGCTGAGAGTTCGTGACCCGCCCGCGAGTCACTCTCGGCAATGACAGAGCTGTACAAAG 720
Db 3350 CTGCTGAGAGTTCGTGACCCGCCCGCGAGTCACTCTCGGCAATGACAGAGCTGTACAAAG 720
QY 721 CTGAGCCATGAGCTTCCGCCCGGAGGTGAGG 751
Db 3410 AGCGGCCGCGACTCTAGAGTGCAGCTTGACG 3440

RESULT 12
US-10-716-725-2
; Sequence 2, Application US/10716725
; Publication No. US2004007613A1
; GENERAL INFORMATION:
; APPLICANT: MAZARAKIS, NICHOLAS
; APPLICANT: AZZOUZ, MINOUN
; APPLICANT: KINGSMAN, SUSAN
; TITLE OF INVENTION: VECTOR SYSTEM
; FILE REFERENCE: 674523-2017.1
; CURRENT APPLICATION NUMBER: US/10/716,725
; PRIOR FILING DATE: 2003-11-19
; PRIOR APPLICATION NUMBER: 10/429,608
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: PCT/GB03/00426
; PRIOR FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: GB 0223076.1
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: GB 0228314.1
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: GB 0318213.6
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: PCT/GB01/04866
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: GB 0122238.9
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: GB 0102339.9
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: GB 0026943.1
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleotide construct pONYBg sequence
US-10-716-725-2

Query Match 84.6%; Score 719; DB 17; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAACCGGGATGATGCCCATCTGTGAGCTGAC 60
Db 2690 ATGTGAGCAAGGCGAGAGCTGTTCAACCGGGATGATGCCCATCTGTGAGCTGAC 2749
QY 61 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGTTCGGCGAGAGGCGATGCTACCTAC 120
Db 2750 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGTTCGGCGAGAGGCGATGCTACCTAC 2809
QY 121 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGTTCGGCGAGAGGCGATGCTACCTAC 180
Db 2810 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGTTCGGCGAGAGGCGATGCTACCTAC 2869
QY 181 CTGCTGAGACCCCTGAGCTTCAATCTGACACACCGGAGAGCTCCGCTGCGCCAC 240
Db 2870 CTGCTGAGACCCCTGAGCTTCAATCTGACACACCGGAGAGCTCCGCTGCGCCAC 2929
QY 241 CAGCAGACATTTCTTCAAGTCCGCCATGCCCGAAGGCTACAGAGAGCGACATCTTC 300
Db 2930 CAGCAGACATTTCTTCAAGTCCGCCATGCCCGAAGGCTACAGAGAGCGACATCTTC 2989
QY 301 TTCAAGAGACAGGCGCAACTACAAAGACCGCGCGCGAGGTGAAGTTGAGAGGCGAC 360
Db 2990 TTCAAGAGACAGGCGCAACTACAAAGACCGCGCGCGAGGTGAAGTTGAGAGGCGAC 3049
QY 361 GTGAACCGCATGAGCTGAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGAGG 420
Db 3050 GTGAACCGCATGAGCTGAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGAGG 3109
QY 421 AAGCTGAGTACACTACACAGCAGCAACGTCATATCATGGCCGACAGAGAAAGAAC 480
Db 3110 AAGCTGAGTACACTACACAGCAGCAACGTCATATCATGGCCGACAGAGAAAGAAC 3169
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGCC 540
Db 3170 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGCC 3229
QY 541 GACCACTACACAGAGAACACCCCATCGGAGACGGCCGCTGCTGCTGCCGACAAACAC 600
Db 3230 GACCACTACACAGAGAACACCCCATCGGAGACGGCCGCTGCTGCTGCCGACAAACAC 3289
QY 601 TACCTGAGACCCAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 660
Db 3290 TACCTGAGACCCAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 3349
QY 661 CTGCTGAGAGTTCGTGACCCGCCCGCGAGTCACTCTCGGCAATGACAGAGCTGTACAAAG 720
Db 3350 CTGCTGAGAGTTCGTGACCCGCCCGCGAGTCACTCTCGGCAATGACAGAGCTGTACAAAG 720
QY 721 CTGAGCCATGAGCTTCCGCCCGGAGGTGAGG 751
Db 3410 AGCGGCCGCGACTCTAGAGTGCAGCTTGACG 3440

RESULT 13
US-09-927-876-98
; Sequence 98, Application US/09927876
; Publication No. US2004005554A1
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabil
; APPLICANT: Campbell, Robert K
; APPLICANT: Kelton, Christie A
; APPLICANT: He, Chaomei
; TITLE OF INVENTION: No. US2004005554A1 Glycoproteins and Methods of Use Thereof
; FILE REFERENCE: 20993-003
; CURRENT APPLICATION NUMBER: US/09/927,876
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/225,035
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/202,724
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98

QY 721 CTTAGCCATGGCT 733
Db 1171 AGCGCCGGCGACT 1183

RESULT 15

US-10-360-149-98
; Sequence 98, Application US/10360149
; Publication No. US20030219786A1
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabli
; APPLICANT: Campbell, Robert K
; APPLICANT: Kelton, Christie A
; APPLICANT: He, Chaomei
; TITLE OF INVENTION: No. US20030219786A1 Glycoproteins and Methods of Use Thereof
; FILE REFERENCE: 20993-003
; CURRENT APPLICATION NUMBER: US/10/360,149
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: US/09/927,876
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/225,035
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/202,724
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 1190
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-360-149-98

Query Match 84.5%; Score 718.6; DB 16; Length 1190;
Best Local Similarity 98.8%; Pred. No. 4.4e-159;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
Db 451 ATGTGAGCAAGGCGAGAGCTGTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 510
QY 61 GCGGAGCTGAAAGGCGCAAGTTTCAAGCTGTCCGCGAGAGGCGAGGCGGCACTTAC 120
Db 511 GCGGAGCTGAAAGGCGCAAGTTTCAAGCTGTCCGCGAGAGGCGAGGCGGCACTTAC 570
QY 121 GCGAAGCTGAGCTGAGTTTCAATCTGACACACCGGCAAGCTGCCGTGCGCCCAACC 180
Db 571 GCGAAGCTGAGCTGAGTTTCAATCTGACACACCGGCAAGCTGCCGTGCGCCCAACC 630
QY 181 CTGTGACCACTTGAAGCTGAGGCTGCAAGTCTTCAAGCTGACCCGCAACCAATGAAG 240
Db 631 CTGTGACCACTTGAAGCTGAGGCTGCAAGTCTTCAAGCTGACCCGCAACCAATGAAG 690
QY 241 CAGCAGCACTTCTTCAAGTCCGCAATGCGGCAAGCTGAGGCTGAGGCTGAGGCTGAG 300
Db 691 CAGCAGCACTTCTTCAAGTCCGCAATGCGGCAAGCTGAGGCTGAGGCTGAGGCTGAG 750
QY 301 TTCAAGGCAAGGCGCAAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
Db 751 TTCAAGGCAAGGCGCAAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 810
QY 361 GTGAACCGGATGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
Db 811 GTGAACCGGATGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 870
QY 421 AAGCTGAGTACAACTGAGCAAGGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 480
Db 871 AAGCTGAGTACAACTGAGCAAGGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 930
QY 481 GGCATCAAGGTGAGCTTCAAGTCCGCAACATCGAGGAGGCGGAGGCTGAGGCTGAGGCTGAG 540

Db 931 GGCATCAAGGTGAGCTTCAAGTCCGCAACATCGAGGAGGCGGAGGCTGAGGCTGAGGCTGAG 990
QY 541 GACCACTACGAGCAAGACACCCCGATGCGGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 600
Db 991 GACCACTACGAGCAAGACACCCCGATGCGGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1050
QY 601 TACCTGAGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 660
Db 1051 TACCTGAGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1110
QY 661 CTGCTGAGGCTTGTGACCGGCGGCGGATCACTTGTGAGTACGAGGCTGAGGCTGAGGCTGAGGCTGAG 720
Db 1111 CTGCTGAGGCTTGTGACCGGCGGCGGATCACTTGTGAGTACGAGGCTGAGGCTGAGGCTGAGGCTGAG 1170
QY 721 CTTAGCCATGGCT 733
Db 1171 AGCGCCGGCGACT 1183

Search completed: May 15, 2004, 06:32:39
Job time : 470 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 03:23:12 ; Search time 3573 Seconds
(without alignments)

8370.613 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atgtgtgacgaagcgagga.....tagatccatgtgtatgc 850

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents NA Main:*

```
1: /cgn2_6/ptodata/2/pna/US0931A_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US0931B_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US0931C_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US0931D_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US0931E_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US0931F_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US0931G_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US0931H_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US0931I_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US0931J_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US0931K_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US0931L_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US0931M_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US0931N_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US0931O_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US0931P_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US0931Q_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US0931R_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US0931S_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US0931T_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US0931U_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US0931V_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US0931W_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US0931X_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US0931Y_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US0931Z_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US0932A_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US0932B_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US0932C_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US0932D_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US0932E_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US0932F_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US0932G_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US0932H_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US0932I_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US0932J_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US0932K_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US0932L_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US0932M_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US0932N_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US0932O_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US0932P_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US0932Q_COMB.seq:*
```

```
44: /cgn2_6/ptodata/2/pna/US1008A_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US1008B_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US1008C_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US1008D_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US1008E_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US1008F_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US1008G_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US1008H_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US1008I_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US1008J_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US1008K_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US1008L_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US1008M_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US1008N_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US1008O_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US1008P_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US1008Q_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US1008R_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US1008S_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US1008T_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US1008U_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US1008V_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US1008W_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US1008X_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US1008Y_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US1008Z_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US1009A_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US1009B_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US1009C_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US1009D_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US1009E_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US1009F_COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US1009G_COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US1009H_COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US1009I_COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US1009J_COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US1009K_COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US1009L_COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US1009M_COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US1009N_COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US1009O_COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US1009P_COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US1009Q_COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US1009R_COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US1009S_COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US1009T_COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US1009U_COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US1009V_COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US1009W_COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US1009X_COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US1009Y_COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US1009Z_COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US1010A_COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US1010B_COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US1010C_COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US1010D_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US1010E_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US1010F_COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US1010G_COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US1010H_COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US1010I_COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US1010J_COMB.seq:*
106: /cgn2_6/ptodata/2/pna/US1010K_COMB.seq:*
107: /cgn2_6/ptodata/2/pna/US1010L_COMB.seq:*
108: /cgn2_6/ptodata/2/pna/US1010M_COMB.seq:*
109: /cgn2_6/ptodata/2/pna/US1010N_COMB.seq:*
110: /cgn2_6/ptodata/2/pna/US1010O_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

```
; LENGTH: 850
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
```

RESULT 2
US-09-191-233-2
; Sequence 2, Application US/09191233A
; GENERAL INFORMATION:

QY 541 GACCACTACCAAGAGAACACCCCATTCGCGACGCGCCCGTGTCTGCTGCCCAACCAAC 600
Db 541 GACCACTACCAAGAGAACACCCCATTCGCGACGCGCCCGTGTCTGCTGCCCAACCAAC 600
QY 601 TACCTGAGCAACCAAGACCGCCCTGAGCAAAAGACCAAGAGAGGCGCATCATGTGTC 660
Db 601 TACCTGAGCAACCAAGACCGCCCTGAGCAAAAGACCAAGAGAGGCGCATCATGTGTC 660
QY 661 CTGCTGAGGTTGTCGACCGCCCGCGGATCACTTCGCGATGAGACGAGTGTACAAGAAG 720
Db 661 CTGCTGAGGTTGTCGACCGCCCGCGGATCACTTCGCGATGAGACGAGTGTACAAGAAG 720
QY 721 CTTAAGCATGCTCTCCCGCCGAGGTGTGAGACAGAGATGATGACGCTGCCCATGTCT 780
Db 721 CTTAAGCATGCTCTCCCGCCGAGGTGTGAGACAGAGATGATGACGCTGCCCATGTCT 780
QY 781 TGTGCCAGAGAGAGGAGATGACCGCTCAACCTGACGCTGTCTGCTGATGATCAAT 840
Db 781 TGTGCCAGAGAGAGGAGATGACCGCTCAACCTGACGCTGTCTGCTGATGATCAAT 840
QY 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 4
US-09-365-089-2

Sequence 2, Application US/09365089
GENERAL INFORMATION:
APPLICANT: Kain, Steve
APPLICANT: Li, Xiangliang
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
FILE REFERENCE: D6100CIP/D1
CURRENT APPLICATION NUMBER: US/09/365,089
CURRENT FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: US 09/191,233
EARLIER FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
US-09-365-089-2

Query Match 100.0%; Score 850; DB 19; Length 850;
Best Local Similarity 100.0%; Pred. No. 2.5e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGAGTGTGCCATCTGTGAGCTGAGC 60
Db 1 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGAGTGTGCCATCTGTGAGCTGAGC 60
QY 61 GGCACAGTAAACGGCCACAAGTTCAAGCTGTCCGCGAGGGCGAGGGGATGCCACTTAC 120
Db 61 GGCACAGTAAACGGCCACAAGTTCAAGCTGTCCGCGAGGGCGAGGGGATGCCACTTAC 120
QY 121 GGCAGCTGACCTGAGAGTTGATCATGACCAACCGGCAAGTGTGCCCTGTGCCCCAAC 180
Db 121 GGCAGCTGACCTGAGAGTTGATCATGACCAACCGGCAAGTGTGCCCTGTGCCCCAAC 180
QY 181 CTGTGACCAACCTGACCTTAAGGCGTGTGACCGCTACCCCGACCATGATGAAG 240
Db 181 CTGTGACCAACCTGACCTTAAGGCGTGTGACCGCTACCCCGACCATGATGAAG 240
QY 241 CAGCAGCACTTCTCAAGTCCGCGATGCGGAGGCTACGTCAGAGGCGGACCATCTTC 300
Db 241 CAGCAGCACTTCTCAAGTCCGCGATGCGGAGGCTACGTCAGAGGCGGACCATCTTC 300

QY 301 TTCAAGAGCAGCGCACTACAGACCCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 360
Db 301 TTCAAGAGCAGCGCACTACAGACCCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 360
QY 361 GTGAACCGCATTCGAGCTTAAGGGCATGCACTTCAAGAGAGCAGCAACATCTGTGGGAC 420
Db 361 GTGAACCGCATTCGAGCTTAAGGGCATGCACTTCAAGAGAGCAGCAACATCTGTGGGAC 420
QY 421 AAGCTGAGTACAACTACAGACCGCAACGCTTATCATGCGCCGACCAAGCAAGAAC 480
Db 421 AAGCTGAGTACAACTACAGACCGCAACGCTTATCATGCGCCGACCAAGCAAGAAC 480
QY 481 GGCATCAAGTGAACCTTCAAGATCCGCAACATCGAGACGCGAGGCTGAGCTGCC 540
Db 481 GGCATCAAGTGAACCTTCAAGATCCGCAACATCGAGACGCGAGGCTGAGCTGCC 540
QY 541 GACCACTACCAAGAGAACACCCCATCGGCGAGCGCCCGTGTCTGCTGCCCAACAC 600
Db 541 GACCACTACCAAGAGAACACCCCATCGGCGAGCGCCCGTGTCTGCTGCCCAACAC 600
QY 601 TACCTGAGCAACCAAGTCCGCGCTGAGCAAGACCCCAACGAGAGCGGATCAATGTC 660
Db 601 TACCTGAGCAACCAAGTCCGCGCTGAGCAAGACCCCAACGAGAGCGGATCAATGTC 660
QY 661 CTGCTGAGGTTGTCGACCGCCCGGATCACTCTCGCATGAGCAGGCTGTACAAGAG 720
Db 661 CTGCTGAGGTTGTCGACCGCCCGGATCACTCTCGCATGAGCAGGCTGTACAAGAG 720
QY 721 CTTAAGCATGCTCTCCCGCCGAGGTGTGAGAGAGAGATGATGACGAGCTGCCATGTCT 780
Db 721 CTTAAGCATGCTCTCCCGCCGAGGTGTGAGAGAGAGATGATGACGAGCTGCCATGTCT 780
QY 781 TGTGCCAGAGAGAGGAGATGACCGTCAACCTGACGCTGTCTGCTGATGATCAAT 840
Db 781 TGTGCCAGAGAGAGGAGATGACCGTCAACCTGACGCTGTCTGCTGATGATCAAT 840
QY 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 5
US-09-931-232-2

Sequence 2, Application US/09931232
GENERAL INFORMATION:
APPLICANT: Kain, Steve
APPLICANT: Li, Xiangliang
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
FILE REFERENCE: CLON075CON
CURRENT APPLICATION NUMBER: US/09/931,232
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/364,946
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/191,233
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 09/062,102
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 60/060,855
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
US-09-931-232-2

Query Match 100.0%; Score 850; DB 38; Length 850;
Best Local Similarity 100.0%; Pred. No. 2.5e-134;

Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGAGCTGTTCACCGGGGTGTGCCCATCTGTGAGCTGGAC 60
Db 1 ATGTGAGCAAGGCGAGAGAGCTGTTCACCGGGGTGTGCCCATCTGTGAGCTGGAC 60
QY 61 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 120
Db 61 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 120
QY 121 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 180
Db 121 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 180
QY 181 CTGTGACCAACCTTGAAGTTCATCTGACCAACCGGCAAGTCCCGTCCCGGCGAGG 240
Db 181 CTGTGACCAACCTTGAAGTTCATCTGACCAACCGGCAAGTCCCGTCCCGGCGAGG 240
QY 241 CAGCAAGCTTCTTCAAGTTCGCGCATGCCCGCAAGGCTTACGTCAGGAGGCGACATCTTC 300
Db 241 CAGCAAGCTTCTTCAAGTTCGCGCATGCCCGCAAGGCTTACGTCAGGAGGCGACATCTTC 300
QY 301 TTCAAGGAGCAAGGCGCAAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACCTTCG 360
Db 301 TTCAAGGAGCAAGGCGCAAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACCTTCG 360
QY 361 GTGAACCGCATGAGCTGAAAGGCGCATGCACTTCAAGGAGGCGAGGCGACATCTTCGAGG 420
Db 361 GTGAACCGCATGAGCTGAAAGGCGCATGCACTTCAAGGAGGCGAGGCGACATCTTCGAGG 420
QY 421 AAGCTGAGTCACTTCAAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGG 480
Db 421 AAGCTGAGTCACTTCAAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGG 480
QY 481 GGCATCAAGTGAAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGG 540
Db 481 GGCATCAAGTGAAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGG 540
QY 541 GACCACTTCAAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 600
Db 541 GACCACTTCAAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 600
QY 601 TACCTGAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 660
Db 601 TACCTGAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 660
QY 661 CTGTGAGGAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 720
Db 661 CTGTGAGGAGTTCAGGAGGCGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 720
QY 721 CTTAGGAGGAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 780
Db 721 CTTAGGAGGAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 780
QY 781 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 781 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 GTGTAGAGTTC 850
Db 841 GTGTAGAGTTC 850

RESULT 6
PCT-US02-17452-87
; Sequence 87, Application PC/TUS0217452
; GENERAL INFORMATION:
; APPLICANT: CHROMOS MOLECULAR SYSTEMS, INC.
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine

APPLICANT: Fleming, Elena
APPLICANT: Stewart, Sandra
APPLICANT: Shellard, Joan
TITLE OR INVENTION: CHROMOSOME-BASED PLATFORMS
FILE REFERENCE: 24601-420PC
CURRENT APPLICATION NUMBER: PCT/US02/17452
CURRENT FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/294,758
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 60/366,891
PRIOR FILING DATE: 2002-03-21
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 87
LENGTH: 4862
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pd2egFP-1N plasmid from Clontech
PCT-US02-17452-87

Query Match 100.0%; Score 850; DB 1; Length 4862;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGAGCTGTTCACCGGGGTGTGCCCATCTGTGAGCTGGAC 60
Db 679 ATGTGAGCAAGGCGAGAGAGCTGTTCACCGGGGTGTGCCCATCTGTGAGCTGGAC 738
QY 61 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 120
Db 739 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 798
QY 121 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 180
Db 799 GGGCACTTAAAGGCGCAAGATTGACGGTGTCCGGCGAGGCGAGGCGAGTCCACTTAC 858
QY 181 CTGTGACCAACCTTGAAGTTCATCTGACCAACCGGCAAGTCCCGTCCCGGCGAGG 240
Db 859 CTGTGACCAACCTTGAAGTTCATCTGACCAACCGGCAAGTCCCGTCCCGGCGAGG 918
QY 241 CAGCAAGCTTCTTCAAGTTCGCGCATGCCCGCAAGGCTTACGTCAGGAGGCGACATCTTC 300
Db 919 CAGCAAGCTTCTTCAAGTTCGCGCATGCCCGCAAGGCTTACGTCAGGAGGCGACATCTTC 978
QY 301 TTCAAGGAGCAAGGCGCAAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTC 360
Db 979 TTCAAGGAGCAAGGCGCAAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTC 1038
QY 361 GTGAACCGCATGAGCTGAAAGGCGCATGCACTTCAAGGAGGCGAGGCGACATCTTCGAGG 420
Db 1039 GTGAACCGCATGAGCTGAAAGGCGCATGCACTTCAAGGAGGCGAGGCGACATCTTCGAGG 480
QY 421 AAGCTGAGTCACTTCAAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 540
Db 1099 AAGCTGAGTCACTTCAAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 1158
QY 481 GGCATCAAGTGAAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 600
Db 1159 GGCATCAAGTGAAGTTCAGGAGGCGCGCGAGGAGTTCAGGAGGCGACATCTTCGAGGAGG 1218
QY 541 GACCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1278
Db 1219 GACCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1338
QY 601 TACCTGAGGAGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 1279 TACCTGAGGAGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1338
QY 661 CTGTGAGGAGTTCAGGAGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 1339 CTGTGAGGAGTTCAGGAGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1398

QY 721 CTTAGCCATGCTTCCCGCGAGAGTGAGAGCAGATGATGCAAGCTGCCATGCT 780
DB 1399 CTTAGCCATGCTTCCCGCGAGAGTGAGAGCAGATGATGCAAGCTGCCATGCT 1458
QY 781 TGTGCCAGAGAGAGGAGAGTGAGAGCTGAGAGCTGCTGCTTCTCTGAGTCAAT 840
DB 1459 TGTGCCAGAGAGAGGAGAGTGAGAGCTGAGAGCTGCTGCTTCTCTGAGTCAAT 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 7
US-10-161-403-87

/ Sequence 87, Application US/10161403
/ GENERAL INFORMATION:
/ APPLICANT: Perkins, Edward
/ APPLICANT: Perez, Carl
/ APPLICANT: Lindenbaum, Michael
/ APPLICANT: Greene, Amy
/ APPLICANT: Leung, Josephine
/ APPLICANT: Fleming, Elena
/ APPLICANT: Stewart, Sandra
/ APPLICANT: Shellard, Joan
/ TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
/ FILE REFERENCE: 24601-420
/ CURRENT APPLICATION NUMBER: US/10/161,403
/ CURRENT FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/294,758
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: 60/366,891
/ PRIOR FILING DATE: 2002-03-21
/ NUMBER OF SEQ ID NOS: 129
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 87
/ LENGTH: 4862
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: pd2egfp-IN plasmid from Clontech
US-10-161-403-87

Query Match 100.0%; Score 850; DB 46; Length 4862;
Best Local Similarity 100.0%; Pred. No. 2,2e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAACGGGGTGTGCCCCATCTGTGAGCTGAC 60
DB 679 ATGTGAGCAAGGCGAGAGCTGTTCAACGGGGTGTGCCCCATCTGTGAGCTGAC 738
QY 61 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 120
DB 739 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 798
QY 121 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 180
DB 799 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 858
QY 181 CTGCTGACCAACCTGACCTACGAGCGTGTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 240
DB 859 CTGCTGACCAACCTGACCTACGAGCGTGTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 918
QY 241 CAGACGACCTTCTTCAAGTCCGATGCGCCGAGAGCTACGTCAGAGAGCGACCATCTTC 300
DB 919 CAGACGACCTTCTTCAAGTCCGATGCGCCGAGAGCTACGTCAGAGAGCGACCATCTTC 978
QY 301 TTCAAGACGACGAGCACTTCAAGAGCCGCGCGAGAGTGAAGTTGAGAGCGAGACCTTC 360
DB 979 TTCAAGACGACGAGCACTTCAAGAGCCGCGCGAGAGTGAAGTTGAGAGCGAGACCTTC 1038
QY 361 GTGAACCGCATCGAGCTGAAGGGGATGACTTCAAGAGAGAGCGGCAATCTGTGGGCGAC 420

DB 1039 GTGAACCGCATCGAGCTGAAGGGGATGACTTCAAGAGAGAGCGGCAATCTGTGGGCGAC 1098
QY 421 AAGCTGAGATCACTTACAAAGCCCAAGCTGTATATCATGAGCCGACAGAGAGAG 480
DB 1099 AAGCTGAGATCACTTACAAAGCCCAAGCTGTATATCATGAGCCGACAGAGAGAG 1158
QY 481 GGCATCAAGTGAAGCTTCAAGATCCGACCAAGATCCGAGAGCGGAGCGTCAAGTCCG 540
DB 1159 GGCATCAAGTGAAGCTTCAAGATCCGACCAAGATCCGAGAGCGGAGCGTCAAGTCCG 1218
QY 541 GACCATCAAGGAG 600
DB 1219 GACCATCAAGGAG 1278
QY 601 TACCTGAG 660
DB 1279 TACCTGAG 1338
QY 661 CTGCTGAG 720
DB 1339 CTGCTGAG 1398
QY 721 CTTAGCCATGCTTCCCGCGAGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 1399 CTTAGCCATGCTTCCCGCGAGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458
QY 781 TGTGCCAG 840
DB 1459 TGTGCCAG 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 8
PCT-US99-28451-12

/ Sequence 12, Application PC/TUS9928451
/ GENERAL INFORMATION:
/ APPLICANT: Li, Xiangdang
/ APPLICANT: Zhao, Xiaoning
/ APPLICANT: Fang, Yu
/ APPLICANT: Jiang, Xin
/ APPLICANT: Duong, Tommy
/ APPLICANT: Kain, Steve R.
/ TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
/ FILE REFERENCE: D6170PCT
/ CURRENT APPLICATION NUMBER: PCT/US99/28451
/ EARLIER FILING DATE: 1999-12-01
/ EARLIER APPLICATION NUMBER: US 09/206,887
/ EARLIER FILING DATE: 1998-12-08
/ NUMBER OF SEQ ID NOS: 14
/ SEQ ID NO 12
/ LENGTH: 4163
/ TYPE: DNA
/ ORGANISM: artificial sequence
/ FEATURE:
/ OTHER INFORMATION: synthesized pNF-(B-d2EGFP construct
PCT-US99-28451-12

Query Match 99.5%; Score 846; DB 1; Length 4163;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAACGGGGTGTGCCCCATCTGTGAGCTGAC 60
DB 241 ATGTGAGCAAGGCGAGAGCTGTTCAACGGGGTGTGCCCCATCTGTGAGCTGAC 300
QY 61 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 120
DB 301 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 360
QY 121 GGGCAGCTAAACGGCCCAAGTTGAGCGTGTCCGGCGAGAGGGCGAGTGCACCTAC 180

Db 361 GGCAGAGTGAACCTTCAAGTTTCTGCAACACCGCAAGTGCCTGCGCCACAC 420
Qy 181 CTGCGACACCCCTGACCTTAAGGCGGTGAGTCTTCAAGCCGTACCCGACCACTGAG 240
Db 421 CTGCGACACCCCTGACCTTAAGGCGGTGAGTCTTCAAGCCGTACCCGACCACTGAG 480
Qy 241 CAGCAGCATTTCTTCAAGTCCGCAATGCGCCGAGGCTACGTCAAGAGCGCACTTTC 300
Db 481 CAGCAGCATTTCTTCAAGTCCGCAATGCGCCGAGGCTACGTCAAGAGCGCACTTTC 540
Qy 301 TTCAAGAGCAGACGCACTCAAGACCCGCGCGAGGTAAGTTGAGGCGCAACCTTG 360
Db 541 TTCAAGAGCAGACGCACTCAAGACCCGCGCGAGGTAAGTTGAGGCGCAACCTTG 600
Qy 361 GTGAACCGCATGAGCTGAGGCGATCGACTTCAAGAGAGAGCGCAATCTTGGGGAC 420
Db 601 GTGAACCGCATGAGCTGAGGCGATCGACTTCAAGAGAGAGCGCAATCTTGGGGAC 660
Qy 421 AAGCTGAGTCAACTAACAAGCCAGCAACGTCTATCATGCGCGACAGAGAGAG 480
Db 661 AAGCTGAGTCAACTAACAAGCCAGCAACGTCTATCATGCGCGACAGAGAGAG 720
Qy 481 GGCATCAAGGTGAATTCAAGATCCGCAACAATCGAGAGCGGAGGTCAGTCCG 540
Db 721 GGCATCAAGGTGAATTCAAGATCCGCAACAATCGAGAGCGGAGGTCAGTCCG 780
Qy 541 GACCACTACAGACGAAACACCCCATGCGGAGAGCGGCGCGGCTGCTGCGCAACAC 600
Db 781 GACCACTACAGACGAAACACCCCATGCGGAGAGCGGCGCGGCTGCTGCGCAACAC 840
Qy 601 TACCTGAGCAACCAAGTCCGCTGAGCAAAAGCCCAAGAGCGGATCACTGTC 660
Db 841 TACCTGAGCAACCAAGTCCGCTGAGCAAAAGCCCAAGAGCGGATCACTGTC 900
Qy 661 CTGCTGAGTTCGTAACCGCGCGCGGATCACTCTGCGCATGAGAGAGCTTGAAGAAG 720
Db 901 CTGCTGAGTTCGTAACCGCGCGCGGATCACTCTGCGCATGAGAGAGCTTGAAGAAG 960
Qy 721 CTGAGCATGAGCTTCCGCGCGAGAGTGAAGAGAGAGTGAAGAGCGTCCCATGCT 780
Db 961 CTGAGCATGAGCTTCCGCGCGAGAGTGAAGAGAGAGTGAAGAGCGTCCCATGCT 1020
Qy 781 TGTGCCAGAGAGAGCGGATGAGACCTGACCTGAGAGCTTGTCTGCTAGATCAAT 840
Db 1021 TGTGCCAGAGAGAGCGGATGAGACCTGACCTGAGAGCTTGTCTGCTAGATCAAT 1080
Qy 841 GTGTAG 846
Db 1081 GTGTAG 1086

RESULT 9
US-09-206-887-12

; Sequence 12, Application US/09206887A
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangqiang
; APPLICANT: Zhao, Xiaoning
; APPLICANT: Pang, Yu
; APPLICANT: Jiang, Xin
; APPLICANT: Duong, Tommy
; APPLICANT: Kahn, Steve R.
; TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
; FILE REFERENCE: D6170
; CURRENT APPLICATION NUMBER: US/09/206,887A
; CURRENT FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 4163
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: synthesized pNF-(B-d2EGFP construct

US-09-206-887-12

Query Match 99.5%; Score 846; DB 16; Length 4163;
Best Local Similarity 100.0%; Pred. No. 1,1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGTGAAGCAAGGCGAGAGCTGTCAACCGGAGTGGTCCCATCTGTGCACTGAC 60
Db 241 ATGGTGAAGCAAGGCGAGAGCTGTCAACCGGAGTGGTCCCATCTGTGCACTGAC 300
Qy 61 GGCAGCGTAAACGGGCAACAAATTCAAGCTGTCCGCGAGAGGCGAGATGCCACTTAC 120
Db 301 GGCAGCGTAAACGGGCAACAAATTCAAGCTGTCCGCGAGAGGCGAGATGCCACTTAC 360
Qy 121 GGCAGAGTCAACCTGAAATTCTATCTGACCAACCGGCAAGCTGCCGTGCTGACCAAC 180
Db 361 GGCAGAGTCAACCTGAAATTCTATCTGACCAACCGGCAAGCTGCCGTGCTGACCAAC 420
Qy 181 CTGCTGACCAACCTGACCTGACGAGGCTGCAAGTCTTCAAGCGCTACCCCGACCATGAG 240
Db 421 CTGCTGACCAACCTGACCTGACGAGGCTGCAAGTCTTCAAGCGCTACCCCGACCATGAG 480
Qy 241 CAGCAGCACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCAAGAGCGGACCATCTTC 300
Db 481 CAGCAGCACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCAAGAGCGGACCATCTTC 540
Qy 301 TTCAAGAGCAGACGCGCACTCAAGACCGCGCGAGGTAAGTTGAGGCGGACACCTTG 360
Db 541 TTCAAGAGCAGACGCGCACTCAAGACCGCGCGAGGTAAGTTGAGGCGGACACCTTG 600
Qy 361 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGAGAGAGCGGCAATCTTGGGGAC 420
Db 601 GTGAACCGCATGAGCTGAGGCGATGCACTTCAAGAGAGAGCGGCAATCTTGGGGAC 660
Qy 421 AAGCTGAGTCAACTAACAAGCCAGCAACGTCTATCATGCGCGCAAGAGAGAGAG 480
Db 661 AAGCTGAGTCAACTAACAAGCCAGCAACGTCTATCATGCGCGCAAGAGAGAGAG 720
Qy 481 GGCATCAAGGTGAATTCAAGATCCGCAACAATCGAGAGCGGAGGTCAGTCCG 540
Db 721 GGCATCAAGGTGAATTCAAGATCCGCAACAATCGAGAGCGGAGGTCAGTCCG 780
Qy 541 GACCACTACAGACGAAACACCCCATGCGGAGAGCGGCGCGGCTGCTGCGCGCAACAC 600
Db 781 GACCACTACAGACGAAACACCCCATGCGGAGAGCGGCGCGGCTGCTGCGCGCAACAC 840
Qy 601 TACCTGAGCAACCAAGTCCGCTGAGCAAAAGCCCAAGAGCGGATCACTGTC 660
Db 841 TACCTGAGCAACCAAGTCCGCTGAGCAAAAGCCCAAGAGCGGATCACTGTC 900
Qy 661 CTGCTGAGTTCGTAACCGCGCGGATCACTCTGCGCATGAGAGAGCTTGAAGAAG 720
Db 901 CTGCTGAGTTCGTAACCGCGCGGATCACTCTGCGCATGAGAGAGCTTGAAGAAG 960
Qy 721 CTGAGCATGAGCTTCCGCGCGAGAGTGAAGAGAGAGTGAAGAGCGTCCCATGCT 780
Db 961 CTGAGCATGAGCTTCCGCGCGAGAGTGAAGAGAGAGTGAAGAGCGTCCCATGCT 1020
Qy 781 TGTGCCAGAGAGAGCGGATGAGACCTGACCTGAGAGCTTGTCTGCTAGATCAAT 840
Db 1021 TGTGCCAGAGAGAGCGGATGAGACCTGACCTGAGAGCTTGTCTGCTAGATCAAT 1080
Qy 841 GTGTAG 846
Db 1081 GTGTAG 1086

RESULT 10
PCT-US99-28451-13

; Sequence 13, Application PCT/US9928451
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangqiang
; APPLICANT: Zhao, Xiaoning


```
APPLICANT: Fang, Yu
APPLICANT: Jiang, Xin
APPLICANT: Duong, Tommy
APPLICANT: Kain, Steve R.
TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
FILE REFERENCE: D6170PCT
CURRENT APPLICATION NUMBER: PCT/US99/28451
EARLIER FILING DATE: 1999-12-01
EARLIER APPLICATION NUMBER: US 09/206,887
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 13
LENGTH: 4300
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthesized pCRE5-d2EGFP
PCT-US99-28451-13
```

Query Match 99.5%; Score 846; DB 1; Length 4300;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 ATGCTAGCAAGGCGGCAAGCTTCAACCGGCGTGTGCCATCTGTGAGCTGAC 60
378 ATGTGAGCAAGGCGGCAAGCTTCAACCGGCGTGTGCCATCTGTGAGCTGAC 437
61 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 120
438 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 497
121 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 180
498 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 557
181 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 240
558 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 617
241 CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGGAGGCGGCGGCAATGCACTTCTTC 300
618 CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGGAGGCGGCGGCAATGCACTTCTTC 677
301 TTCAGGACGACGCGGCACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 360
678 TTCAGGACGACGCGGCACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 737
361 GTGAACCGCATGAGCTGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 420
738 GTGAACCGCATGAGCTGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 797
421 AAGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 480
798 AAGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 857
481 GGCATCAAGGTGAATTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 540
858 GGCATCAAGGTGAATTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 917
541 GACCACTTACAGGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 600
918 GACCACTTACAGGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 977
601 TACCTGAGCAACCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 660
978 TACCTGAGCAACCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 1037
661 CTGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 720
1038 CTGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 1097
721 CTGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 780
```

```
1098 CTTAGCCATGAGCTTCCCGCGGAGGTGAGGAGCAGATGATGCAACGCTGCCATGCTCT 1157
781 TGTGCGCCAGGAGGCGGAGTGAACCGTCAACCTTGAGGCGTGTCTTGTAGATCATAT 840
1158 TGTGCGCCAGGAGGCGGAGTGAACCGTCAACCTTGAGGCGTGTCTTGTAGATCATAT 1217
841 GTGTAG 846
1218 GTGTAG 1223
```

RESULT 11
US-09-206-887-13
Sequence 13, Application US/09206887A

```
GENERAL INFORMATION:
APPLICANT: Li, Xiangliang
APPLICANT: Zhao, Xiaoning
APPLICANT: Fang, Yu
APPLICANT: Jiang, Xin
APPLICANT: Duong, Tommy
APPLICANT: Kain, Steve R.
TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
FILE REFERENCE: D6170
CURRENT APPLICATION NUMBER: US/09/206,887A
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 13
LENGTH: 4300
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthesized pCRE5-d2EGFP
US-09-206-887-13
```

Query Match 99.5%; Score 846; DB 16; Length 4300;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 ATGCTAGCAAGGCGGCAAGCTTCAACCGGCGTGTGCCATCTGTGAGCTGAC 60
378 ATGTGAGCAAGGCGGCAAGCTTCAACCGGCGTGTGCCATCTGTGAGCTGAC 437
61 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 120
438 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 497
121 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 180
498 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 557
181 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 240
558 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 617
498 GCGCAGTAAACGGCCACAGTTACGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 557
181 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 240
558 CTCTGACCAACCTTGAACCTTACGCGGCGTGTCCGGGAGGCGGCGGCAATGCACTTAC 617
618 CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGGAGGCGGCGGCAATGCACTTCTTC 677
301 TTCAGGACGACGCGGCACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 360
618 CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGGAGGCGGCGGCAATGCACTTCTTC 677
301 TTCAGGACGACGCGGCACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 360
678 TTCAGGACGACGCGGCACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 737
361 GTGAACCGCATGAGCTGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 420
738 GTGAACCGCATGAGCTGAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 797
421 AAGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 480
798 AAGCTGAGTAACTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 857
481 GGCATCAAGGTGAATTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 540
858 GGCATCAAGGTGAATTTCAAGGCGGCGGCGGAGGCGGCGGCAATGCACTTCTTC 917
```

QY	541	GACCACTTACAGAGAAACACCCCATTCGGAGACGACCCGCTGTCTGCCGACCAACAC	600
Db	918	GACCATTACAGACAAACACCCCATTCGGAGACGACCCGCTGTCTGCCGACCAACAC	977
QY	601	TACCTGAGCACCAGTCGCGCCCTGAGCAAAAGACCCACGAGAGCGCGATCATATGTC	660
Db	978	TACCTGAGCACCAGTCGCGCCCTGAGCAAAAGACCCACGAGAGCGCGATCATATGTC	1037
QY	661	CTGCTGGAGTTCGTGACCGCGCCGCGGAGATCATCTTCGGCATTGACGAGCTGTCAAGAG	720
Db	1038	CTGCTGGAGTTCGTGACCGCGCCGCGGAGATCATCTTCGGCATTGACGAGCTGTCAAGAG	1097
QY	721	CTTACCATGAGCTTCCCGCCGAGGCTGTGAGAGACAGATATGCGACGCTGCCATGTCT	780
Db	1098	CTTACCATGAGCTTCCCGCCGAGGCTGTGAGAGACAGATATGCGACGCTGCCATGTCT	1157
QY	781	TGTGGCCAGAGAGCGGGATGTGACCGCTCACCTTCGACGCTGTGCTTCTGTAGATTAAT	840
Db	1158	TGTGGCCAGAGAGCGGGATGTGACCGCTCACCTTCGACGCTGTGCTTCTGTAGATTAAT	1217
QY	841	GTGTAG 846	
Db	1218	GTGTAG 1223	

```

RESULT 12
US-10-3332-733-21
; Sequence 21, Application US/10332733
; GENERAL INFORMATION:
; APPLICANT: Margarette Odenthal and Diana Jung
; TITLE OF INVENTION: Gene Expression And Reporter Expression
; TITLE OF INVENTION: In Myofibroblasts And Myofibroblast-like Cells
; FILE REFERENCE: 1472/66806
; CURRENT APPLICATION NUMBER: US/10/3332, 733
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 845
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence: EGFP
; US-10-3332-733-21

```

Query Match	99.4%	Score 845	DB 49	Length 845
Best Local Similarity	100.0%	Pred. No. 1.8e-133		
Matches 845	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	ATGTGAGCAAGGGCAGAGAGCTGTTTCAACCGGGGTGGAGCCATCTGGTCAAGTGGAC	60
Dp	1	ATGTGAGCAAGGGCAGAGAGCTGTTTCAACCGGGGTGGAGCCATCTGGTCAAGTGGAC	60
QY	61	GGGACGCTAAAGGCGCCAAAGTTCAAGCGGTTCGGGCGAGGGCGAGGGCGATGCCACTTAC	120
Dp	61	GGGACGCTAAAGGCGCCAAAGTTCAAGCGGTTCGGGCGAGGGCGAGGGCGATGCCACTTAC	120
QY	121	GGCAAGCTACCCCTGAAGTTCAATCTGACCAACAACGGGCAAGCTGACCGCTGGGCCACC	180
Dp	121	GGCAAGCTACCCCTGAAGTTCAATCTGACCAACAACGGGCAAGCTGACCGCTGGGCCACC	180
QY	121	GGCAAGCTACCCCTGAAGTTCAATCTGACCAACAACGGGCAAGCTGACCGCTGGGCCACC	180
Dp	121	GGCAAGCTACCCCTGAAGTTCAATCTGACCAACAACGGGCAAGCTGACCGCTGGGCCACC	180
QY	181	CTGTGACCAACCTGACCTTACGGGCGTGCAGTGTCTTCAAGCGGCTTACCCGACCAATGAAG	240
Dp	181	CTGTGACCAACCTGACCTTACGGGCGTGCAGTGTCTTCAAGCGGCTTACCCGACCAATGAAG	240
QY	241	CAGCAGCACTTCTTCAAGTCCGCGATGCGCCGAGAGGCTACGTCCAGAGAGCCACATCTTC	300
Dp	241	CAGCAGCACTTCTTCAAGTCCGCGATGCGCCGAGAGGCTACGTCCAGAGAGCCACATCTTC	300
QY	301	TTTCAAGAGAGAGGCGCACTTCAAGACCCGGCCGAGAGGTTCAGAGGCGACACCTTC	360
Dp	301	TTTCAAGAGAGAGGCGCACTTCAAGACCCGGCCGAGAGGTTCAGAGGCGACACCTTC	360

QY	361	GTGGAACCCGATCGAGCTGAAGGCAATCGACTTCAAGAGAGACGGGCAACATCCGCGGGCAC	420
Db	361	GTAACCCGATGAGCTGAAGGCAATCGACTTCAAGAGAGACGGGCAACATCCGCGGGCAC	420
QY	421	AAGCTGAGTACAACTACAAACGACCAACGTCATATCATGGCCGACAAAGCAGAAGAC	480
Db	421	AAGCTGAGTACAACTACAAAGCCCAACGTCATATCATGGCCGACAAAGCAGAAGAC	480
QY	481	GGCATCAAGTGAATCTTCAAGATCCGGCCCAACATCTGAGAGACGGGACGGCTGCAATCCGC	540
Db	481	GGCATCAAGGTAATCTTCAAGATCCGGCCCAACATCTGAGAGACGGGACGGCTGCAATCCGC	540
QY	541	GACCACTACACAGCAACAACCCCAATCGGCGACGGGCCCGTGTGCTGTCGCCGCAACAC	600
Db	541	GACCACTACACAGCAACAACCCCAATCGGCGACGGGCCCGTGTGCTGTCGCCGCAACAC	600
QY	601	TACCTGAGCACCCAGTCCGGCCCTTGAGCAAAAGCCCAACGAAGAGGGCATCATATGATC	660
Db	601	TACCTGAGCACCCAGTCCGGCCCTTGAGCAAAAGCCCAACGAAGAGGGCATCATATGATC	660
QY	661	CTGCTGAGATTGATGACCGCGCGCGCGAGATCATCTCTCGCATGAGACGAGCTGTACAAAGA	720
Db	661	CTGCTGAGATTGATGACCGCGCGCGCGAGATCATCTCTCGCATGAGACGAGCTGTACAAAGA	720
QY	721	CTTAGCCATGGCTTCCCGCCGAGGTGGAGAGACAGATGATGAGCAAGCTGGCCCATGTCT	780
Db	721	CTTAGCCATGGCTTCCCGCCGAGGTGGAGAGACAGATGATGAGCAAGCTGGCCCATGTCT	780
QY	781	TGTGGCCAGAGAGCGGATGGAACGTCACACTGACAGCCTTGCTTCTCTAGATCAAT	840
Db	781	TGTGGCCAGAGAGCGGATGGAACGTCACACTGACAGCCTTGCTTCTCTAGATCAAT	840
QY	841	GTGTA 845	
Db	841	GTGTA 845	

```

RESULT 13
PCT-US03-28111-3
Sequence 3, Application PC/TUS0328111
GENERAL INFORMATION:
APPLICANT: Bear, et al.
TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of
FILE REFERENCE: Use Inhereof
CURRENT APPLICATION NUMBER: PCT/US03/28111
CURRENT FILING DATE: 2003-09-05
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 6748
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Plasmid
PCT-US03-28111-3

```

Query Match	84.9%	Score 721.8;	DB 1;	Length 6748;
Best Local Similarity	99.0%;	Pred. No. 1.1e-112;		
Matches 726;	Conservative	0;	Mismatches 7;	Indels 0;
				Gaps 0;

Oy 1 ATGTGTGACGAGGGCGAGGAGCTGTTCACCGGGGGTGGTCCATCTTGAGCTGGAC 60
 Db 2778 ATGTGTGACGAGGGCGAGGAGCTGTTCACCGGGGGTGGTCCATCTTGAGCTGGAC 2833
 Oy 61 GGCACGCTAAACGGCCCAAGTTACAGCGTGTCCGGCGAGGGCGAGTGCCTTAC 120
 Db 2838 GGCACGCTAAACGGCCCAAGTTACAGCGTGTCCGGCGAGGGCGAGTGCCTTAC 2893
 Oy 121 GGCAGCTGACCTGAAATTCACTGCACACCGGCAAGCTGCCCTGCGCCAC 180

US-10-152-040-28

Query Match 84.8%; Score 721.2; DB 46; Length 10417;
Best Local Similarity 97.6%; Pred. No. 1.3e-112;
Matches 732; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
QY 1 ATGTTGAGCAAGGCGAGGAGCTGTTACCGGGGTGTGCGCATCTGTGTGAGCTGAGC 60
DB 743 ATGTTGAGCAAGGCGAGGAGCTGTTACCGGGGTGTGCGCATCTGTGTGAGCTGAGC 802
QY 61 GGGGAGCTAAAGGCGCAAGTTCAAGCTGTCCGGGAGGAGGCGCATGCCACTTAC 120
DB 803 GGGGAGCTAAAGGCGCAAGTTCAAGCTGTCCGGGAGGAGGCGCATGCCACTTAC 862
QY 121 GCGAAGCTGACCTGAAGTTTCATCTGCAACCGGCAAGCTGCGCCCTGCGCCACC 180
DB 863 GCGAAGCTGACCTGAAGTTTCATCTGCAACCGGCAAGCTGCGCCCTGCGCCACC 922
QY 181 CTCTGACACCACTGACCTTACCGGCTGCAAGTGTCTTACCGCTTACCGGCAATGAA 240
DB 923 CTCTGACACCACTGACCTTACCGGCTGCAAGTGTCTTACCGCTTACCGGCAATGAA 982
QY 241 CAGCAAGATTCTTCAAGTCCGCGCATGCGCGAAGGTAAGTCCAGAGGCGCATCTTC 300
DB 983 CAGCAAGATTCTTCAAGTCCGCGCATGCGCGAAGGTAAGTCCAGAGGCGCATCTTC 1042
QY 301 TTCAAGAGAGGCGCACTACAAAGACCGCGCGGAGTGAAGTTGAGGGCGACACCTG 360
DB 1043 TTCAAGAGAGGCGCACTACAAAGACCGCGCGGAGTGAAGTTGAGGGCGACACCTG 1102
QY 361 GTGAACCGCATGAGCTGAAGGGCATTCGACTTCAAGAGAGGCGCAACATCTTGGGGC 420
DB 1103 GTGAACCGCATGAGCTGAAGGGCATTCGACTTCAAGAGAGGCGCAACATCTTGGGGC 1162
QY 421 AAGCTGAGATACACTACAGCAAGCCCAAGTCTATATGATGCGCGCAAGAGAAAG 480
DB 1163 AAGCTGAGATACACTACAGCAAGCCCAAGTCTATATGATGCGCGCAAGAGAAAG 1222
QY 481 GGCATCAAGTGAATCTTCAAGATCCGCCCAACATGAGAGCGGCAAGCTGAGCTGCC 540
DB 1223 GGCATCAAGTGAATCTTCAAGATCCGCCCAACATGAGAGCGGCAAGCTGAGCTGCC 1282
QY 541 GACCACTACAGAGAAACACCCCATGCGGAGCGGCCCGTGTCTGCTCGACAAACAC 600
DB 1283 GACCACTACAGAGAAACACCCCATGCGGAGCGGCCCGTGTCTGCTCGACAAACAC 1342
QY 601 TACCTGAGGACCAAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGCATCAATGCTC 660
DB 1343 TACCTGAGGACCAAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGCATCAATGCTC 1402
QY 661 CTGCTGAGATGCTGACCGCGCGGATCACTCTGCGATGAGCGAGCTGTACAAGAG 720
DB 1403 CTGCTGAGATGCTGACCGCGCGGATCACTCTGCGATGAGCGAGCTGTACAAGAG 1462
QY 721 CTTAGCCATGGCTTCCCGCGGAGGTGAG 750
DB 1463 CTTAGCCATGGCTTCCCGCGGAGGTGAG 1492
```

Search completed: May 15, 2004, 06:21:47
Job time : 3577 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 03:30:36 / Search time 170 Seconds
(without alignments)
5764.986 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atgtgtgagcaagcgagcgagga.....tagatcaatgtgtatgacgc 850

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 1660298 seqs, 576498597 residues

Total number of hits satisfying chosen parameters: 3320596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents: NA New: *
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	850	100.0	4336	US-10-679-191A-1	Sequence 1, Appl
2	719.2	84.6	14083	US-10-789-400-3	Sequence 3, Appl
3	718.6	84.5	1190	US-10-811-081-98	Sequence 98, Appl
4	718.6	84.5	4944	US-10-826-523-55	Sequence 55, Appl
5	718.6	84.5	4952	US-10-826-523-52	Sequence 52, Appl
6	718.6	84.5	5451	US-10-485-281-6	Sequence 6, Appl
7	718.6	84.5	9013	US-10-826-523-50	Sequence 50, Appl
8	718.6	84.5	5759	US-09-954-483B-14	Sequence 14, Appl
9	717.4	84.4	1704	US-10-488-758-8	Sequence 8, Appl
10	717.4	84.4	5041	PCT-US04-06378-6	Sequence 6, Appl
11	717.4	84.4	5041	US-10-790-455-6	Sequence 6, Appl
12	717.4	84.4	8693	US-10-475-962-54	Sequence 54, Appl
13	717.4	84.4	717	US-10-785-862-9	Sequence 9, Appl
14	717.4	84.4	8591	US-10-672-764A-42	Sequence 42, Appl
15	717.4	84.4	10089	US-10-672-764A-53	Sequence 53, Appl
16	717.4	84.4	38203	US-10-672-764A-55	Sequence 55, Appl
17	717.4	84.4	38203	US-10-672-764A-56	Sequence 56, Appl
18	714	84.0	765	US-10-129-207B-7	Sequence 7, Appl
19	711	83.6	720	US-10-311-030-10	Sequence 10, Appl
20	707.4	83.2	4941	US-10-826-523-53	Sequence 53, Appl
21	707.4	83.2	4943	US-10-826-523-54	Sequence 54, Appl
22	705.8	83.0	9012	US-10-826-523-49	Sequence 49, Appl
23	704.2	82.8	4944	US-10-826-523-56	Sequence 56, Appl
24	704.2	82.8	4951	US-10-826-523-51	Sequence 51, Appl
25	704.2	82.8	8999	US-10-826-523-48	Sequence 48, Appl
26	703	82.7	1722	US-10-831-622-107	Sequence 107, App

27	589	69.3	600	1	PCT-US04-10982-168	Sequence 168, App
28	584.2	68.7	597	1	PCT-US04-10982-36	Sequence 36, Appl
29	584.2	68.7	597	1	PCT-US04-10982-60	Sequence 60, Appl
30	584.2	68.7	597	1	PCT-US04-10982-164	Sequence 164, Appl
31	584.2	68.7	600	1	PCT-US04-10982-38	Sequence 38, Appl
32	584.2	68.7	600	1	PCT-US04-10982-166	Sequence 166, Appl
33	584.2	68.7	600	1	PCT-US04-10982-166	Sequence 68, Appl
34	582.6	68.5	597	1	PCT-US04-10982-84	Sequence 84, Appl
35	582.6	68.5	597	1	PCT-US04-10982-152	Sequence 152, App
36	582.6	68.5	597	1	PCT-US04-10982-70	Sequence 70, Appl
37	582.6	68.5	600	1	PCT-US04-10982-86	Sequence 86, Appl
38	582.6	68.5	600	1	PCT-US04-10982-154	Sequence 154, App
39	582.6	68.5	600	1	PCT-US04-10982-32	Sequence 32, Appl
40	581	68.4	597	1	PCT-US04-10982-88	Sequence 88, Appl
41	581	68.4	597	1	PCT-US04-10982-96	Sequence 96, Appl
42	581	68.4	597	1	PCT-US04-10982-100	Sequence 100, Appl
43	581	68.4	597	1	PCT-US04-10982-160	Sequence 160, App
44	581	68.4	597	1	PCT-US04-10982-174	Sequence 174, App
45	581	68.4	597	1	PCT-US04-10982-174	Sequence 174, App

ALIGNMENTS

RESULT 1
US-10-679-191A-1

Sequence 1, Application US/10679191A

GENERAL INFORMATION:

APPLICANT: Moon, Randall T.

FILE REFERENCE: Dorky, Richard I

CURRENT APPLICATION NUMBER: US/10/679,191A

PRIOR FILING DATE: 2003-10-03

PRIOR APPLICATION NUMBER: US 60/416,504

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 4336

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: plasmid

NAME/KEY: misc feature

LOCATION: (1)..(1)

FEATURE:

NAME/KEY: misc feature

LOCATION: (38)..(243)

OTHER INFORMATION: SV40 polyA

FEATURE:

NAME/KEY: misc feature

LOCATION: (3005)..(3374)

OTHER INFORMATION: Topflash enhancer

FEATURE:

NAME/KEY: misc feature

LOCATION: (3485)..(4327)

OTHER INFORMATION: EGFP+pest

US-10-679-191A-1

Query Match 100.0%; Score 850; DB 6; Length 4336;

Best Local Similarity 100.0%; Pred. No. 4.2e-150;

Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGCGGCGAGGAGCTGTCACCGGGGTGTCCCATCTGTCAGCTGAC 60

Db 3485 ATGTGAGCAAGCGGCGAGGAGCTGTCACCGGGGTGTCCCATCTGTCAGCTGAC 3544

QY 61 GGGGAGCTAAAGGCGCAAGTTGACGCGTGTCCGGGAGGCGAGGCGCATGCGCACTAC 120

Db 3545 GGGGAGCTAAAGGCGCAAGTTGACGCGTGTCCGGGAGGCGAGGCGCATGCGCACTAC 3604

APPLICANT: Campbell, Robert K
APPLICANT: Kelton, Christie A
APPLICANT: He, Chaoemei
TITLE OF INVENTION: Novel Glycoproteins and Methods of Use Thereof
FILE REFERENCE: 20993-003
CURRENT APPLICATION NUMBER: US/10/811,081
CURRENT FILING DATE: 2004-03-25
PRIOR APPLICATION NUMBER: US/10/457,047
PRIOR FILING DATE: 2003-06-05
PRIOR APPLICATION NUMBER: US/10/360,149
PRIOR FILING DATE: 2003-02-06
PRIOR APPLICATION NUMBER: US/09/927,876
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/225,035
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/202,724
PRIOR FILING DATE: 2000-05-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 98
LENGTH: 1190
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-811-081-98

Query Match 84.5%; Score 718.6; DB 6; Length 1190;
Best Local Similarity 98.8%; Pred. No. 1.2e-125;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGTGTTCACCGGGGTGTGCCCCATCTGTGAGCTGAGC 60
DB 451 ATGTGAGCAAGGGGAGAGAGTGTTCACCGGGGTGTGCCCCATCTGTGAGCTGAGC 510
QY 61 GGGCAGCTAAAGGGGCAAGTTCAGCGTGTCCGCGAGGGGAGGGGCGATGCCACTAC 120
DB 511 GGGCAGCTAAAGGGGCAAGTTCAGCGTGTCCGCGAGGGGAGGGGCGATGCCACTAC 570
QY 121 GGCAGCTGAGCCCTGAAGTTCATCTGCACCAACGGGCAAGCTCCGTCGCGCCCAACC 180
DB 571 GGCAGCTGAGCCCTGAAGTTCATCTGCACCAACGGGCAAGCTCCGTCGCGCCCAACC 630
QY 181 CTCTGACCAACCTGACCTTACGGCGTGCAGTGTTCAGCCGCTACCCCGACCAATGAA 240
DB 631 CTCTGACCAACCTGACCTTACGGCGTGCAGTGTTCAGCCGCTACCCCGACCAATGAA 690
QY 241 CAGCAGCACTTCTTCAAGTTCGCGCCATGCGCCGAGGCTTACCTCAGAGAGCGACCTTC 300
DB 691 CAGCAGCACTTCTTCAAGTTCGCGCCATGCGCCGAGGCTTACCTCAGAGAGCGACCTTC 750
QY 301 TTCAAGGAGCAGCGCAACTACAGACCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 360
DB 751 TTCAAGGAGCAGCGCAACTACAGACCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 810
QY 361 GTGAACCGCATGAGCTGAAAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGGGGAC 420
DB 811 GTGAACCGCATGAGCTGAAAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGGGGAC 870
QY 421 AAGCTGAGTCACTTCAATACAGCAGCAACGCTTATATATGCGGAGAGAGAGAGAG 480
DB 871 AAGCTGAGTCACTTCAATACAGCAGCAACGCTTATATATGCGGAGAGAGAGAGAG 930
QY 481 GGCATCAAGGTGAATCTTCAAGATTCGCGCAACAATCGAGAGAGCGAGCGTGCAGCTGCC 540
DB 931 GGCATCAAGGTGAATCTTCAAGATTCGCGCAACAATCGAGAGAGCGAGCGTGCAGCTGCC 990
QY 541 GACCATCAAGCAGCAAGAACCCCATGCGCGAGCGCGCGCGTGTCTGTGCGCGACCAAC 600
DB 991 GACCATCAAGCAGCAAGAACCCCATGCGCGAGCGCGCGCGTGTCTGTGCGCGACCAAC 1050
QY 601 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGAGCCCAAGAGAGAGCGAGATCATGTGTC 660

DB 1051 TACCTGAGCAGCCAGTCCGCGCTGAGCAAGAGCCCAAGAGAGCGAGATCATGTGTC 1110
QY 661 CTGCTGAGTGTGAGTACCGCGCGCGGAGTCACTCTCGGCATGAGCGAGCTGAGAGAG 720
DB 1111 CTGCTGAGTGTGAGTACCGCGCGCGGAGTCACTCTCGGCATGAGCGAGCTGAGAGTAA 1170
QY 721 CTAGCCATGAGCT 733
DB 1171 AGCGCGCGGAGCT 1183

RESULT 4
US-10-826-523-55
Sequence 55, Application US/10826523
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
APPLICANT: BEAM, TERESA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
FILE REFERENCE: 835910-92098
CURRENT APPLICATION NUMBER: US/10/826,523
CURRENT FILING DATE: 2004-04-19
PRIOR APPLICATION NUMBER: US/10/001,189
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 55
LENGTH: 4944
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
US-10-826-523-55

Query Match 84.5%; Score 718.6; DB 6; Length 4944;
Best Local Similarity 98.8%; Pred. No. 1.4e-125;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGTGTTCACCGGGGTGTGCCCCATCTGTGAGCTGAGC 60
DB 1716 ATGTGAGCAAGGGGAGAGAGTGTTCACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1775
QY 61 GGCAGCTAAAGGGGCAAGTTCAGCGTGTCCGCGAGGGGAGGGCGATGCCACTAC 120
DB 1776 GGCAGCTAAAGGGGCAAGTTCAGCGTGTCCGCGAGGGGAGGGCGATGCCACTAC 1835
QY 121 GGCAGCTGAGCCCTGAAGTTCATCTGCACCAACGGGCAAGCTCCGTCGCGCCCAACC 180
DB 1836 GGCAGCTGAGCCCTGAAGTTCATCTGCACCAACGGGCAAGCTCCGTCGCGCCCAACC 1895
QY 181 CTCTGACCAACCTGACCTTACGGCGTGCAGTGTTCAGCCGCTACCCCGACCAATGAA 240
DB 1896 CTCTGACCAACCTGACCTTACGGCGTGCAGTGTTCAGCCGCTACCCCGACCAATGAA 1955
QY 241 CAGCAGCACTTCTTCAAGTTCGCGCCATGCGCCGAGGCTTACCTCAGAGAGCGACCTTC 300
DB 1956 CAGCAGCACTTCTTCAAGTTCGCGCCATGCGCCGAGGCTTACCTCAGAGAGCGACCTTC 2015
QY 301 TTCAAGGAGCAGCGCAACTACAGACCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 360
DB 2016 TTCAAGGAGCAGCGCAACTACAGACCCGCGCGAGTGAAGTTGAGGGCGACACCTTG 2075
QY 361 GTGAACCGCATGAGCTGAAAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGGGGAC 420
DB 2076 GTGAACCGCATGAGCTGAAAGGCGATCGACTTCAAGAGAGAGCGCAACATCTTGGGGAC 2135

QY 421 AAGCTGAGTACAACTACAAAGCCAAAGTCTATATCATGCGCAGCAAGCAAGAAAC 480
DB 2136 AAGCTGAGTACAACTACAAAGCCAAAGTCTATATCATGCGCAGCAAGCAAGAAAC 2195
QY 481 GGGATTAAGTGAACCTTCAAGATCCGCCCAACATGAGAGAGCGGCTGACGCTCGCC 540
DB 2196 GGGATTAAGTGAACCTTCAAGATCCGCCCAACATGAGAGAGCGGCTGACGCTCGCC 2255
QY 541 GACCACTACAGAGAAACCCCATCGGAGACGCGCCGCTGCTGCTGCGCCAGCAACAC 600
DB 2256 GACCACTACAGAGAAACCCCATCGGAGACGCGCCGCTGCTGCTGCGCCAGCAACAC 2315
QY 601 TACCTGAGACACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 660
DB 2316 TACCTGAGACACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 2375
QY 661 CTGCTGAGTGTGTCACCGCCCGCGGATGATCTTCTGCGATGAGCAAGCTGTACAAAG 720
DB 2376 CTGCTGAGTGTGTCACCGCCCGCGGATGATCTTCTGCGATGAGCAAGCTGTACAAAG 2435
QY 721 CTGAGCATGGCT 733
DB 2436 AGCGGCCCGGACT 2448

RESULT 5
US-10-826-523-52/c
; Sequence 52, Application US/10826523
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/826,523
; CURRENT FILING DATE: 2004-04-19
; PRIOR APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 4952
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EGFP
; OTHER INFORMATION: sequence
US-10-826-523-52

Query Match 84.5%; Score 718.6; DB 6; Length 4952;
Best Local Similarity 98.8%; Pred. No. 1.4e-125;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 60
DB 2187 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 2128
QY 61 GGGAGCTAAACGGCGACAAAGTTCAAGCTGTCCGCGAGAGGCGAGGCGATGCACTTAC 120
DB 2127 GGGAGCTAAACGGCGACAAAGTTCAAGCTGTCCGCGAGAGGCGAGGCGATGCACTTAC 2068
QY 121 GGGAGCTAAACCTTGAAGTTCAATCTGACACCAAGCTGCGCTGCTGCTGCGCCACC 180
DB 2067 GGGAGCTAAACCTTGAAGTTCAATCTGACACCAAGCTGCGCTGCTGCTGCGCCACC 2008
QY 181 CTGTGACCAACCTGACCTTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

DB 2007 CTGTGACCAACCTGACCTTACGAGGCTGCAAGTGTTCAGCCGCTACCCGACCATGAG 1948
QY 241 CAGCAAGCTTCTTCAAGTCCGCCATGCGCGAAGGCTTACGAGAGCGACCATCTTC 300
DB 1947 CAGCAAGCTTCTTCAAGTCCGCCATGCGCGAAGGCTTACGAGAGCGACCATCTTC 1888
QY 301 TTCAAGAGCAAGGCACTTCAAGACCCGCGCGAGGTTGAGAGGCGACACCTTG 360
DB 1887 TTCAAGAGCAAGGCACTTCAAGACCCGCGCGAGGTTGAGAGGCGACACCTTG 1828
QY 361 GTGAACCGCATGAGCTGAAAGGCGATGCACTTCAAGAGAGAGCGGTAATCTTGGGGAC 420
DB 1827 GTGAACCGCATGAGCTGAAAGGCGATGCACTTCAAGAGAGAGCGGTAATCTTGGGGAC 1768
QY 421 AAGCTGAGTCAACTTACAAAGCCCAAGCGTATATCATGCGCGCAAGCAAGAAAC 480
DB 1767 AAGCTGAGTCAACTTACAAAGCCCAAGCGTATATCATGCGCGCAAGCAAGAAAC 1708
QY 481 GGCATCAAGTGAACCTTCAAGATCCGCCCAACATGAGAGAGCGAGGCTGACGCTGCGC 540
DB 1707 GGCATCAAGTGAACCTTCAAGATCCGCCCAACATGAGAGAGCGAGGCTGACGCTGCGC 1648
QY 541 GACCACTACAGAGAAACCCCATCGGCGAAGGCGCCCGCTGCTGCTGCGCGAACAAC 600
DB 1647 GACCACTACAGAGAAACCCCATCGGCGAAGGCGCCCGCTGCTGCTGCGCGAACAAC 1588
QY 601 TACCTGAGACACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGCGCATGACATGCTC 660
DB 1587 TACCTGAGACACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGCGCATGACATGCTC 1528
QY 661 CTGCTGAGTGTGTCACCGCCCGCGGATGATCTTCTGCGATGAGCAAGCTGTACAAAG 720
DB 1527 CTGCTGAGTGTGTCACCGCCCGCGGATGATCTTCTGCGATGAGCAAGCTGTACAAAG 1468
QY 721 CTGAGCATGGCT 733
DB 1467 AGCGGCCCGGACT 1455

RESULT 6
US-10-485-281-6/c
; Sequence 6, Application US/10485281
; GENERAL INFORMATION:
; APPLICANT: Sandig, Volker
; APPLICANT: Jordan, Ingo
; TITLE OF INVENTION: RNA Amplification System Using Plant Components in
; TITLE OF INVENTION: Animal Cells
; FILE REFERENCE: 040026us/3H
; CURRENT APPLICATION NUMBER: US/10/485,281
; CURRENT FILING DATE: 2004-01-29
; PRIOR APPLICATION NUMBER: PCT/DE02/02863
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5451
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vector pLuo-79
US-10-485-281-6

Query Match 84.5%; Score 718.6; DB 6; Length 5451;
Best Local Similarity 98.8%; Pred. No. 1.4e-125;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 60
DB 4898 ATGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 4839
QY 61 GGGAGCTAAACGGCGACAAAGTTCAAGCTGTCCGCGAGAGGCGAGGCGATGCACTTAC 120
DB 4838 GGGAGCTAAACGGCGACAAAGTTCAAGCTGTCCGCGAGAGGCGAGGCGATGCACTTAC 4779

QY 121 GGCAAGCTGACCCCTGAGTTTCATCTGACCAACCGGACGCTGCCCTGCTGACCAACC 180
DB 4778 GGCAAGCTGACCCCTGAGTTTCATCTGACCAACCGGACGCTGCCCTGCTGACCAACC 4719
QY 181 CTCGTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 240
DB 4728 CTCGTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 4659
QY 241 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGAGGTACGTCAGGAGGCGCACTCTTC 300
DB 4658 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGAGGTACGTCAGGAGGCGCACTCTTC 4599
QY 301 TTCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
DB 4598 TTCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4539
QY 361 GTGAACCGCATGAGCTGAGAGGCGATCGACTTCAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 4538 GTGAACCGCATGAGCTGAGAGGCGATCGACTTCAGAGAGAGAGAGAGAGAGAGAGAG 4479
QY 421 AAGCTGAGATCAACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 480
DB 4478 AAGCTGAGATCAACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4419
QY 481 GGCATCAAGGTGACCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 4418 GGCATCAAGGTGACCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 4359
QY 541 GACCACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600
DB 4358 GACCACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4299
QY 601 TACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 660
DB 4298 TACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4239
QY 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGAGAGAGAGAGAG 720
DB 4238 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGAGAGAGAGAGAG 4179
QY 721 CTGAGCAGTGGCT 733
DB 4178 AGCGGCGCGGACT 4166

RESULT 7
US-10-826-523-50
; Sequence 50, Application US/10826523
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/826,523
; PRIOR FILING DATE: 2004-04-19
; PRIOR APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 9013
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-Egfp

OTHER INFORMATION: sequence
US-10-826-523-50
Query Match 84.5%; Score 718.6; DB 6; Length 9013;
Best Local Similarity 98.8%; Fred. No. 1.4e-125;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 AAGTGACCAAGGAGGAG 60
DB 7262 AAGTGACCAAGGAGGAG 7321
QY 61 GCGCAGCTGAAACGCGCAGCAAGTTCAAGCTGTCGCGGAGAGAGAGAGAGAGAGAGAGAG 120
DB 7322 GCGCAGCTGAAACGCGCAGCAAGTTCAAGCTGTCGCGGAGAGAGAGAGAGAGAGAGAGAG 7381
QY 121 GCGCAGCTGACCTGTAAGTTCACTGACACACCGGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 7382 GCGCAGCTGACCTGTAAGTTCACTGACACACCGGAGAGAGAGAGAGAGAGAGAGAGAG 7441
QY 181 CTGCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 240
DB 7442 CTGCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 7501
QY 241 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 7502 CAGCAGCACTTCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7561
QY 301 TTCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
DB 7562 TTCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7621
QY 361 GTGAACCGCATGAGCTGAGAGGCGATCGACTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 7622 GTGAACCGCATGAGCTGAGAGGCGATCGACTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 7681
QY 421 AAGCTGAGTCAACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 480
DB 7682 AAGCTGAGTCAACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7741
QY 481 GGCATCAAGGTGACCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 7742 GGCATCAAGGTGACCTTCAAGTCCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 7801
QY 541 GACCACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600
DB 7802 GACCACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7861
QY 601 TACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 660
DB 7862 TACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7921
QY 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGAGAGAGAGAGAGAG 720
DB 7922 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGAGAGAGAGAGAGAG 7981
QY 721 CTGAGCAGTGGCT 733
DB 7982 AGCGGCGCGGACT 7994

RESULT 8
US-09-954-483B-14
; Sequence 14, Application US/09954483B
; GENERAL INFORMATION:
; APPLICANT: Siebel, Christian
; APPLICANT: Brennan, Thomas J.
; TITLE OF INVENTION: Methods of Producing Cells and Animals
; TITLE OF INVENTION: Compiling Targeted Gene Modifications and Compositions
; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: RMES-02
; CURRENT APPLICATION NUMBER: US/09/954,483B
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/232,957

Db 1585 TACCTGAGCACCAGTCCGCCCTGAGCAAAAGACCCCAAGAGCGGATCATGTGTC 1644
Qy 661 CTGCTGAGTTCGTGAGACCGCCCGCGGATCTCTCGGATGAGAGAGCTGTACAA 719
Db 1645 CTGCTGAGTTCGTGAGACCGCCCGCGGATCTCTCGGATGAGAGAGCTGTACAA 1703

RESULT 10
PCT-US04-06378-6
; Sequence 6, Application PC/TUS0406378
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Integrase-mediated Transformation
; FILE REFERENCE: A181 1080.1
; CURRENT APPLICATION NUMBER: PCT/US04/06378
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 5041
; TYPE: DNA
; ORGANISM: Plasmid pCMV-EGFP-attB
PCT-US04-06378-6

Query Match 84.4%; Score 717.4; DB 1; Length 5041;
Best Local Similarity 99.2%; Pred. No. 2,3e-125;
Matches 721; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 60
Db 4301 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 4360
Qy 61 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 120
Db 4361 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4420
Qy 121 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 180
Db 4421 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4480
Qy 181 CTGTGAGCAAGCTGAGCTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 240
Db 4481 CTGTGAGCAAGCTGAGCTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4540
Qy 241 CAGACGACCTTCAAGTCCGATGCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 300
Db 4541 CAGACGACCTTCAAGTCCGATGCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4600
Qy 301 TTCAAGGACGAGCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 360
Db 4601 TTCAAGGACGAGCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4660
Qy 361 GTGAACCGGATCGAGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 420
Db 4661 GTGAACCGGATCGAGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4720
Qy 421 AAGCTGAGTACCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 480
Db 4721 AAGCTGAGTACCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4780
Qy 481 GGCATCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 540
Db 4781 GGCATCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 4840
Qy 541 GACCACTACCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 600
Db 4841 GACCACTACCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 4900
Qy 601 TACCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAG 660
Db 4901 TACCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAG 4960
Qy 661 CTGCTGAGTTCGTGAGACCGCCCGCGGATCTCTCGGATGAGAGAGCTGTACAAAG 720

Db 4961 CTGCTGAGTTCGTGAGACCGCCCGCGGATCTCTCGGATGAGAGAGCTGTACAA 5020
Qy 721 CTGAGCC 727
Db 5021 AGCGGCC 5027

RESULT 11
US-10-790-455-6
; Sequence 6, Application US/10790455
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Integrase-mediated Transformation
; FILE REFERENCE: A181 1080.1
; CURRENT APPLICATION NUMBER: US/10/790,455
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 5041
; TYPE: DNA
; ORGANISM: Plasmid pCMV-EGFP-attB
US-10-790-455-6

Query Match 84.4%; Score 717.4; DB 6; Length 5041;
Best Local Similarity 99.2%; Pred. No. 2,3e-125;
Matches 721; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 60
Db 4301 ATGTGAGCAAGGCGGAGAGCTGTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 4360
Qy 61 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 120
Db 4361 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4420
Qy 121 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 180
Db 4421 GGGACCTAAACGGCCCAAGTTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4480
Qy 181 CTGTGAGCAAGCTGAGCTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 240
Db 4481 CTGTGAGCAAGCTGAGCTGAGGTGTCGGGCGAGGGCGAGGGGATGCCACTAC 4540
Qy 241 CAGACGACCTTCAAGTCCGATGCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 300
Db 4541 CAGACGACCTTCAAGTCCGATGCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4600
Qy 301 TTCAAGGACGAGCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 360
Db 4601 TTCAAGGACGAGCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4660
Qy 361 GTGAACCGGATCGAGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 420
Db 4661 GTGAACCGGATCGAGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4720
Qy 421 AAGCTGAGTACCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 480
Db 4721 AAGCTGAGTACCACTACCAAGACCCGCGAGGCTGAGGTGTCGGGCGAGGGGATGCCACTAC 4780
Qy 481 GGCATCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 540
Db 4781 GGCATCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 4840
Qy 541 GACCACTACCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 600
Db 4841 GACCACTACCAAGGTGAATCTCAAGATCCGCAACAATCGAGAGCGGAGCTGAGCTGCC 4900
Qy 601 TACCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAG 660
Db 4901 TACCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAGCAAGGCTGAG 4960

QY 661 CTGCTGAGTTGTCAGCCGCGGAGTACTCTGCGATGAGAGAGCTGTACAGAG 720
Db 4961 CTGCTGAGTTGTCAGCCGCGGAGTACTCTGCGATGAGAGAGCTGTACAGAG 5020
QY 721 CTTAGCC 727
Db 5021 AGCGGCC 5027

RESULT 12
US-10-475-962-54
; Sequence 54, Application US/10475962
; GENERAL INFORMATION:
; APPLICANT: CHAMON, PIERRE
; APPLICANT: GHISELINCK, NORBERT B.
; APPLICANT: SCHNUTGEN, FRANK
; TITLE OF INVENTION: METHOD FOR THE STABLE INVERSION OF DNA SEQUENCE BY SITE-SPECIFIC
; FILE REFERENCE: 37991-0025
; CURRENT APPLICATION NUMBER: US/10/475,962
; PRIOR FILING DATE: 2003-10-27
; PRIOR APPLICATION NUMBER: PCT/IB02/02493
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 09/843,150
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 54
; LENGTH: 8693
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequence: DNA
; OTHER INFORMATION: sequence of plasmid pLEXR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(360)
; OTHER INFORMATION: SV40 promoter, sense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1050)
; OTHER INFORMATION: loxP1 site, sense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130)
; OTHER INFORMATION: loxS11 site, sense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1170)..(2050)
; OTHER INFORMATION: EGFP polyA gene, sense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2060)..(5700)
; OTHER INFORMATION: lacZ polyA gene, antisense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5710)
; OTHER INFORMATION: loxP1 site, antisense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5790)
; OTHER INFORMATION: loxS11 site, antisense
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5830)..(8693)
; OTHER INFORMATION: vector sequence
US-10-475-962-54

Query Match 84.4%, Score 717.4, DB 6, Length 8693;

Best Local Similarity 99.9%; Pred. No. 2.4e-125;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGGCCATCTCGTGTAGCTGAG 60
Db 1172 ATGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGGCCATCTCGTGTAGCTGAG 1231
QY 61 GGGCAGCTAAACGGCCCAAGTTCAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTAC 120
Db 1232 GGGCAGCTAAACGGCCCAAGTTCAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTAC 1291
QY 121 GGGCAGCTGACCCCTGAGAGTTCATCTGACACACCGGCAAGTGTCCCTGGCCCAACC 180
Db 1292 GGGCAGCTGACCCCTGAGAGTTCATCTGACACACCGGCAAGTGTCCCTGGCCCAACC 1351
QY 181 CTGCTGACCAACCTGACCTACGCGGTGAGTGTCTTACGCGCTACCCCGACCATGTAAG 240
Db 1352 CTGCTGACCAACCTGACCTACGCGGTGAGTGTCTTACGCGCTACCCCGACCATGTAAG 1411
QY 241 CAGCAGCACTTCTTCAAGTCCGCCATGCCGAGGCTACGTCAGAGCGCACCATCTTC 300
Db 1412 CAGCAGCACTTCTTCAAGTCCGCCATGCCGAGGCTACGTCAGAGCGCACCATCTTC 1471
QY 301 TTCAAGAGCAAGCGGCACTACCAAGACCCGGCGCGAGGTGAAGTTCAGAGCGCAACCTTG 360
Db 1472 TTCAAGAGCAAGCGGCACTACCAAGACCCGGCGCGAGGTGAAGTTCAGAGCGCAACCTTG 1531
QY 361 GTGAACCGCATTCAGCTGAGAGGCGGATGACTTCAAGAGAGCGGCAACATCTGAGGCAAC 420
Db 1532 GTGAACCGCATTCAGCTGAGAGGCGGATGACTTCAAGAGAGCGGCAACATCTGAGGCAAC 1531
QY 421 AAGCTGAGTACCACTACCAACAGCCACACGCTTATATCATGCGCGAACAGAGAAC 480
Db 1592 AAGCTGAGTACCACTACCAACAGCCACACGCTTATATCATGCGCGAACAGAGAAC 1651
QY 481 GGCATCAAGGTGAAGCTTCAAGATCCGCCCAACATCATGAGACCGGACGTCAGCTGCC 540
Db 1652 GGCATCAAGGTGAAGCTTCAAGATCCGCCCAACATCATGAGACCGGACGTCAGCTGCC 1711
QY 541 GACCACCTACGACGAGACACCCCATCGGCGAGCGGCGGCTGCGCCGACCAACAC 600
Db 1712 GACCACCTACGACGAGACACCCCATCGGCGAGCGGCGGCTGCGCCGACCAACAC 1771
QY 601 TACCTGACCAACCAAGTCCGCTTGAAGCAAAAGCCCAAGAGAGCGGATCATGATGTC 660
Db 1772 TACCTGACCAACCAAGTCCGCTTGAAGCAAAAGCCCAAGAGAGCGGATCATGATGTC 1831
QY 661 CTGCTGAGTTGTCAGCCGCGGAGTACTCTGCGATGAGAGAGCTGTACAGAG 719
Db 1832 CTGCTGAGTTGTCAGCCGCGGAGTACTCTGCGATGAGAGAGCTGTACAGAG 1890

RESULT 13
US-10-785-862-9
; Sequence 9, Application US/10785862
; GENERAL INFORMATION:
; APPLICANT: Donald Danforth Plant Science Center
; APPLICANT: Fauquet, Claude M.
; APPLICANT: Padmanabhan, Chellappan
; APPLICANT: Ramachandran, Vanitharani
; TITLE OF INVENTION: siRNA-mediated inhibition of gene expression in plant cells
; FILE REFERENCE: C35621/104850
; CURRENT APPLICATION NUMBER: US/10/785,862
; PRIOR FILING DATE: 2004-02-24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GFP coding sequence
US-10-785-862-9

Query Match 84.4%; Score 717; DB 6; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGACCAAGGCGGAGAGAGCTTTCAACCGGGGTGTGCTCCATCTGTGAGCTGGAC 60
DB 1 ATGTGACCAAGGCGGAGAGAGCTTTCAACCGGGGTGTGCTCCATCTGTGAGCTGGAC 60
QY 61 GCGCAGCTAAACCGGCGCAAGATTCCGCGTCCGCGAGGGGAGGGGATGCCCTTAC 120
DB 61 GCGCAGCTAAACCGGCGCAAGATTCCGCGTCCGCGAGGGGAGGGGATGCCCTTAC 120
QY 121 GCGAAGCTGACCTTGAAGTTCAATCTGCAACAACCGGCAAGCTGCCCTGCGCCAC 180
DB 121 GCGAAGCTGACCTTGAAGTTCAATCTGCAACAACCGGCAAGCTGCCCTGCGCCAC 180
QY 181 CTGCTGACCACTTGAAGCTTCAAGGCTGCTTCAAGCCGCTACCCCGACCACTGAAG 240
DB 181 CTGCTGACCACTTGAAGCTTCAAGGCTGCTTCAAGCCGCTACCCCGACCACTGAAG 240
QY 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCGAGGCTACGTCCAGAGGAGGACCATCTTC 300
DB 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCGAGGCTACGTCCAGAGGAGGACCATCTTC 300
QY 301 TTCAAGGACGACGCGCACTACAGAGACCCGCGCGAGTGAAGTTGAGAGGCGACCTTC 360
DB 301 TTCAAGGACGACGCGCACTACAGAGACCCGCGCGAGTGAAGTTGAGAGGCGACCTTC 360
QY 361 GTGAACCGCATCGAGCTGAGAGGCGCATCTTCAAGAGAGAGCGGCAATCTTGGGGAC 420
DB 361 GTGAACCGCATCGAGCTGAGAGGCGCATCTTCAAGAGAGAGCGGCAATCTTGGGGAC 420
QY 421 AAGCTGAGTAACTTCAAGCAAGCCCAAGCTTATCATGCGCGCAAGAGAGAGAC 480
DB 421 AAGCTGAGTAACTTCAAGCAAGCCCAAGCTTATCATGCGCGCAAGAGAGAGAC 480
QY 481 GGCATCAAGGTGAAGCTTCAAGATCCGCGCAACATCGAGAGCGGAGCTGCGCC 540
DB 481 GGCATCAAGGTGAAGCTTCAAGATCCGCGCAACATCGAGAGCGGAGCTGCGCC 540
QY 541 GACCACTAACGAGAGAACCCCGCATCGCGAGCGCGCGCGCTGCTGCCGACCAAC 600
DB 541 GACCACTAACGAGAGAACCCCGCATCGCGAGCGCGCGCGCTGCTGCCGACCAAC 600
QY 601 TACCTGAGCAACCGAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTC 660
DB 601 TACCTGAGCAACCGAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTC 660
QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATCATCTCTCGGCAATGAGAGCTGTACAG 717
DB 661 CTGCTGAGTTCGTGACCGCGCGCGGATCATCTCTCGGCAATGAGAGCTGTACAG 717
```

RESULT 14

```
US-10-672-764A-42
; Sequence 42, Application US/10672764A
; GENERAL INFORMATION:
; APPLICANT: Jolly, Chris
; TITLE OF INVENTION: Immunoglobulin Compositions and Methods
; FILE REFERENCE: 13311.1001U
; CURRENT APPLICATION NUMBER: US/10/672,764A
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 8591
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a laboratory from synthetic material
US-10-672-764A-42
```

Query Match 84.4%; Score 717; DB 6; Length 8591;
Best Local Similarity 100.0%; Pred. No. 2.8e-125;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGACCAAGGCGGAGAGAGCTTTCAACCGGGGTGTGCTCCATCTGTGAGCTGGAC 60
DB 671 ATGTGACCAAGGCGGAGAGAGCTTTCAACCGGGGTGTGCTCCATCTGTGAGCTGGAC 730
QY 61 GCGCAGCTAAACCGGCGCAAGATTCCGCGTCCGCGAGGGGAGGGGATGCCCTTAC 120
DB 731 GCGCAGCTAAACCGGCGCAAGATTCCGCGTCCGCGAGGGGAGGGGATGCCCTTAC 120
QY 121 GCGAAGCTGACCTTGAAGTTCAATCTGCAACAACCGGCAAGCTGCCCTGCGCCAC 180
DB 791 GCGAAGCTGACCTTGAAGTTCAATCTGCAACAACCGGCAAGCTGCCCTGCGCCAC 180
QY 181 CTGCTGACCACTTGAAGCTTCAAGGCTGCTTCAAGCCGCTACCCCGACCACTGAAG 240
DB 851 CTGCTGACCACTTGAAGCTTCAAGGCTGCTTCAAGCCGCTACCCCGACCACTGAAG 240
QY 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCGAGGCTACGTCCAGAGGAGGACCATCTTC 300
DB 911 CAGACGACTTCTTCAAGTCCGCGCATGCGCGAGGCTACGTCCAGAGGAGGACCATCTTC 300
QY 301 TTCAAGGACGACGCGCACTACAGAGACCCGCGCGAGTGAAGTTGAGAGGCGACCTTC 360
DB 971 TTCAAGGACGACGCGCACTACAGAGACCCGCGCGAGTGAAGTTGAGAGGCGACCTTC 360
QY 361 GTGAACCGCATCGAGCTGAGAGGCGCATCTTCAAGAGAGAGCGGCAATCTTGGGGAC 420
DB 1031 GTGAACCGCATCGAGCTGAGAGGCGCATCTTCAAGAGAGAGCGGCAATCTTGGGGAC 420
QY 421 AAGCTGAGTAACTTCAAGCAAGCCCAAGCTTATCATGCGCGCAAGAGAGAGAC 480
DB 1091 AAGCTGAGTAACTTCAAGCAAGCCCAAGCTTATCATGCGCGCAAGAGAGAGAC 480
QY 481 GGCATCAAGGTGAAGCTTCAAGATCCGCGCAACATCGAGAGCGGAGCTGCGCC 540
DB 1151 GGCATCAAGGTGAAGCTTCAAGATCCGCGCAACATCGAGAGCGGAGCTGCGCC 540
QY 541 GACCACTAACGAGAGAACCCCGCATCGCGAGCGCGCGCGCTGCTGCCGACCAAC 600
DB 1211 GACCACTAACGAGAGAACCCCGCATCGCGAGCGCGCGCGCTGCTGCCGACCAAC 600
QY 601 TACCTGAGCAACCGAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTC 660
DB 1271 TACCTGAGCAACCGAGTCCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTC 660
QY 661 CTGCTGAGTTCGTGACCGCGCGCGGATCATCTCTCGGCAATGAGAGCTGTACAG 717
DB 1331 CTGCTGAGTTCGTGACCGCGCGCGGATCATCTCTCGGCAATGAGAGCTGTACAG 717
```

RESULT 15

```
US-10-672-764A-53
; Sequence 53, Application US/10672764A
; GENERAL INFORMATION:
; APPLICANT: Jolly, Chris
; TITLE OF INVENTION: Immunoglobulin Compositions and Methods
; FILE REFERENCE: 13311.1001U
; CURRENT APPLICATION NUMBER: US/10/672,764A
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 10089
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a laboratory from synthetic material
; NAME/KEY: misc_feature
; LOCATION: (1)...(10089)
```

OTHER INFORMATION: n = A,T,C or G
US-10-672-764A-53

Query Match 84.4%; Score 717; DB 6; Length 10089;
Best Local Similarity 100.0%; Pred. No. 2.9e-125;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTGAGCAAGGCGAGAGCTTTACCGGGGTGTGCTCCATCTGTGTGAGCTGAGC	60
DB	1029	ATGCTGAGCAAGGCGAGAGCTTTACCGGGGTGTGCTCCATCTGTGTGAGCTGAGC	1088
QY	61	GGCGACGTAAAGCGCCACAAGTTTCAGCGTCCGGGCGAGGGGAGGCGATGCCCTTAC	120
DB	1089	GGCGACGTAAAGCGCCACAAGTTTCAGCGTCCGGGCGAGGGGAGGCGATGCCCTTAC	1148
QY	121	GGCAAGCTGACCTTGAAGTTTCACTGACACACCGGCAAGCTGCCCTGTGCGCCACG	180
DB	1149	GGCAAGCTGACCTTGAAGTTTCACTGACACACCGGCAAGCTGCCCTGTGCGCCACG	1208
QY	181	CTGCTGACCACTTCACTTACCTTACGCGGTGCACTGCTTCACTGCTTCACTGCTTCA	240
DB	1209	CTGCTGACCACTTCACTTACCTTACGCGGTGCACTGCTTCACTGCTTCACTGCTTCA	1268
QY	241	CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTTACGTCGAGGAGGCACTTTC	300
DB	1269	CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTTACGTCGAGGAGGCACTTTC	1328
QY	301	TTCAAGGACGAGCGGCACTTCAAGACCGCGCGCGAGGTGAAGTTTCAAGGCGGAC	360
DB	1329	TTCAAGGACGAGCGGCACTTCAAGACCGCGCGCGAGGTGAAGTTTCAAGGCGGAC	1388
QY	361	GTGAACGCGCATCGAGCTGAGGAGGATGCACTTCAAGGAGGACGCGACATCTTGGGCGAC	420
DB	1389	GTGAACGCGCATCGAGCTGAGGAGGATGCACTTCAAGGAGGACGCGACATCTTGGGCGAC	1448
QY	421	AAGCTGAGTAACTAACCACTACAGCCACAAGCTTATATCATGCGCGACCAAGCAGAGAAC	480
DB	1449	AAGCTGAGTAACTAACCACTACAGCCACAAGCTTATATCATGCGCGACCAAGCAGAGAAC	1508
QY	481	GGCATCAAGGTGAAGCTTCAAGATCCGCCACAACATCGAGAGCGGCAAGCTGCGC	540
DB	1509	GGCATCAAGGTGAAGCTTCAAGATCCGCCACAACATCGAGAGCGGCAAGCTGCGC	1568
QY	541	GACCACTAACGAGGAGACACCCCGCATGCGCGACGCGCCCGCTGTGCTGCGCACAC	600
DB	1569	GACCACTAACGAGGAGACACCCCGCATGCGCGACGCGCCCGCTGTGCTGCGCACAC	1628
QY	601	TACCTGAGCAACCGATCGCCCTGAGCAAGACCCCAAGAGGAGCGGATCACATGATC	660
DB	1629	TACCTGAGCAACCGATCGCCCTGAGCAAGACCCCAAGAGGAGCGGATCACATGATC	1688
QY	661	CTGCTGAGATTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAG	717
DB	1689	CTGCTGAGATTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAG	1745

Search completed: May 15, 2004, 06:24:44
Job time : 173 secs